



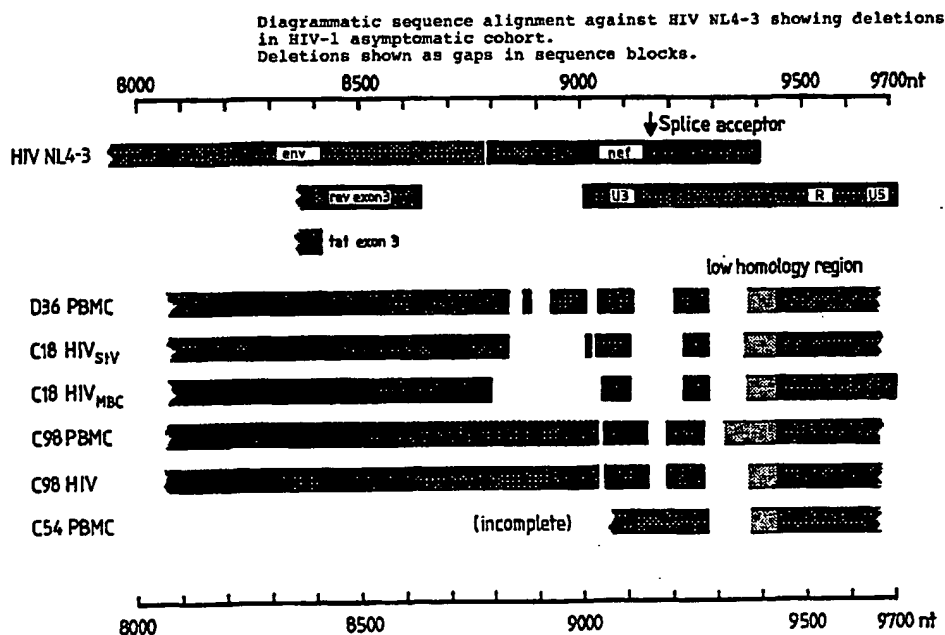
PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 7/02, 7/04, 15/48, C12Q 1/68, 1/70, A61K 39/21</b>		A1	(11) International Publication Number: <b>WO 95/21912</b>
			(43) International Publication Date: 17 August 1995 (17.08.95)
(21) International Application Number: <b>PCT/AU95/00063</b>		(74) Agents: HUGHES, E. John, L. et al.; Davies Collison Cave, 1 Little Collins Street, Melbourne, VIC 3000 (AU).	
(22) International Filing Date: 14 February 1995 (14.02.95)			
(30) Priority Data: PM 3864 14 February 1994 (14.02.94) AU PM 4002 21 February 1994 (21.02.94) AU PN 0284 23 December 1994 (23.12.94) AU		(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).	
(71) Applicants (for all designated States except US): THE MACFARLANE BURNET CENTRE FOR MEDICAL RESEARCH LIMITED [AU/AU]; Yarra Bend Road, Fairfield, VIC 3078 (AU). AUSTRALIAN RED CROSS SOCIETY (NSW DIVISION) [AU/AU]; 153 Clarence Street, Sydney, NSW 2000 (AU).		Published With international search report.	
(72) Inventors; and (75) Inventors/Applicants (for US only): DEACON, Nicholas, John [AU/AU]; 51 Elliott Avenue, Balwyn, VIC 3105 (AU). LEARMONT, Jennifer, Catherine [AU/AU]; 88 Patongo Street, Patongo, NSW 2256 (AU). MCPHEE, Dale, Alan [AU/AU]; 14 Freeman Street, Fitzroy, VIC 3065 (AU). CROWE, Suzanne [AU/AU]; 1 Palermo Street, South Yarra, VIC 3141 (AU). COOPER, David [AU/AU]; 7 Carrington Avenue, Bellvue Hill, NSW 2023 (AU).			

(54) Title: NON-PATHOGENIC STRAINS OF HIV-1



(57) Abstract

The present invention relates to non-pathogenic strains of HIV-1 and to components, parts, fragments and derivatives thereof and to genetic sequences derived therefrom and their use in the development of diagnostic and therapeutic compositions for the treatment and prophylaxis of AIDS and AIDS-related disorders. The present invention also relates to a method for attenuating pathogenic strains of HIV-1 by mutagenizing particular regions of the HIV-1 genome.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Larvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## NON-PATHOGENIC STRAINS OF HIV-1

5 The present invention relates to non-pathogenic strains of HIV-1 and to components, parts, fragments and derivatives thereof and to genetic sequences derived therefrom and their use in the development of diagnostic and therapeutic compositions for the treatment and prophylaxis of AIDS and AIDS-related disorders. The present invention also relates to a method for attenuating pathogenic strains of HIV-1 by mutagenizing particular  
10 regions of the HIV-1 genome.

Bibliographic details of the publications referred to in this specification are collected at the end of the description. Sequence Identity Numbers (SEQ ID NOs.) for the nucleotide and amino acid sequences referred to in the specification are defined  
15 following the bibliography.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but  
20 not the exclusion of any other element or integer or group of elements or integers.

Genomic nucleotide sequences of HIV-1 strains referred to herein are represented by their corresponding DNA sequence.

25 Exemplary viral isolates referred to herein as "C18" and "C98" were deposited at the PHLS Centre for Applied Microbiology and Research, European Collection of Animal Cell Cultures (ECACC), Division of Biologies, Porton Down, Salisbury, Wiltshire SP4 OJG. C18 was deposited on 17 October, 1994 under Provisional Accession Number V94101706 and C98 was deposited on 31 October, 1994 under Provisional Accession  
30 Number V941031169. Viral isolate "C54" was deposited at ECACC on 14 February, 1995 under Provisional Accession Number \_\_\_\_\_.

A summary of particular deletion mutants of HIV-1 of the present invention referred to herein is given in Figure 11.

5 Acquired Immune Deficiency Syndrome (AIDS) and AIDS related disorders are the clinical result of infection by Human Immunodeficiency Virus type I (HIV-1) (Barre-Sinoussi *et al*, 1983). Infection by HIV-1 is generally characterised by progressive immune system damage (Teeuwssen *et al*, 1990; Clerici *et al*, 1989) leading to opportunistic infections, malignancies or wasting syndrome that constitute clinically-defined AIDS (Busch *et al*, 1991; Klaslow *et al*, 1990).

10

The high mortality rate of individuals infected with HIV-1 together with the social and economic consequences of the continuing HIV-1 epidemic has created an urgent need for a safe and effective treatment and/or prophylaxis against the devastating effects of AIDS. However, despite over a decade of high level scientific research into the  
15 pathogenesis of HIV-1 and the clinical manifestations of the disease, together with a detailed molecular analysis of the virus, there has been little success in the development of an effective vaccine. To date, the most effective therapy is treatment with zidovudine (AZT) which delays the onset of full blown AIDS and alleviates to some extent the symptoms of HIV-1 infection. However, AZT is not an innocuous compound and AZT,  
20 metabolic products thereof or impurities therein can cause a number of side effects which limit long term treatment with the drug. Furthermore, AZT resistant isolates have been reported during treatment. Clearly, therefore, a need exists to develop alternative strategies in preventing and treating HIV-1 infection.

25 The initial phases of HIV-1 infection are summarised by Levy (1993) as involving attachment, fusion and nucleocapsid entry. These phases have been the traditional foci in research into development of antiviral strategies. The molecular events at the virus genomic level have also been the subject of intense scientific research with an aim being the development of a live attenuated vaccine as a possible approach for the treatment or  
30 prophylaxis of HIV-1 infection.



There is a high variable rate of progression from initial HIV-1 infection to AIDS which reflects a rapidly changing pathogen and variable immune response of the host to infection (Sheppard *et al*, 1993). With regards to the latter, HIV-1 can be considered  
5 as a heterogenous group of viruses differing at the genetic level with concomitant variable pathogenicity. For example, HIV-1 strains can differ in their capacity to kill cells. Furthermore, it appears that HIV-1 strains evolve in a host after infection and that the evolution varies depending on the tissues infected by the virus. The major sites in the genome apparently responsible for biological and pathological variation are the  
10 highly variable envelope region (Cheng-Mayer *et al*, 1991; Shioda *et al*, 1992; Hwang, *et al* 1991; Sullivan *et al*, 1993; Groenink *et al*, 1993) and the viral regulatory regions such as *tat* (Leguern *et al*, 1993). The genetic complexity of the HIV-1 group of viruses together with their variable pathogenicity, are major difficulties in the development of live vaccines, genetic vaccines or component vaccines.

15 Notwithstanding the highly pathogenic nature of HIV-1, there are some reports of long term survival of subjects infected with the virus (Learmont *et al*, 1992; Levy, 1993; Sheppard *et al*, 1993; Lifson *et al* 1991). It is not always clear, however, whether a benign course following HIV-1 infection is due to host factors, viral factors or other  
20 unknown factors. There are reports that most infected people have at least laboratory evidence of progressive immune system damage in the form of CD4+ cell loss (Lang *et al*, 1989) and defective immune responses (Clerici *et al*, 1989).

Although simian monkeys have been used as an *in vivo* model for HIV and Simian  
25 Immunodeficiency Virus (SIV) infection, a major handicap in AIDS research has been the absence of suitable *in vivo* models to study the pathogenesis of the disease and, in particular, to study the viruses involved in benign infection. The need for a suitable *in vivo* model is heightened by the fact that results obtained *in vitro* cannot necessarily be extrapolated to what occurs *in vivo*. This was clearly observed by Mosier *et al* (1993)  
30 where conflicting results were obtained in animals compared to cell cultures.

Despite the absence of suitable *in vivo* models, considerable scientific research has been directed to attenuating HIV-1 strains by mutagenesis of the virus genome. Deletions in the *nef* gene have been implicated in attenuated strains of SIV and their use in providing protective effects in monkeys (Daniel *et al*, 1992). However, there are conflicting reports on the possible negative influence the *nef* gene product has on the rate or extent of virus replication (Terwilliger *et al*, 1986; Luciw *et al*, 1987; Niederman *et al*, 1989; Kim *et al*, 1989; Hammes *et al*, 1989). In fact, Kim *et al* (1989) found that *nef* did not affect HIV-1 replication or HIV-1 long terminal repeat (LTR)-driven CAT expression. Kestler III *et al* (1991) found that the *nef* gene is required for full pathogenic potential in SIV. However, such is the complexity of the HIV-1 group of viruses and the variability of immune responses between individuals let alone different species, that it is far from clear whether *nef* deleted strains of HIV-1 would behave similarly to *nef* deleted strains of SIV-I. There is a need, therefore, in order to investigate the possibility of *nef* deleted HIV-1 strain as a vaccine candidate, to identify individuals infected with such modified viruses.

Learmont *et al* (1992) reported that a cohort of five persons infected with blood products from a single HIV-1 infected donor have remained asymptomatic from up to about 10-14 years after infection. Subsequently, a sixth person has been identified as being part of the cohort. Both the donor and recipients were HIV-1 seropositive but with no indications of clinical symptoms of HIV-1 related disease and CD4+ cell number and  $\beta_2$ -microglobulin levels have remained in the normal range. The identification of this cohort of benignly infected individuals provides a unique *in vivo* model in which the pathogenesis of HIV-1 infection can be studied at the clinical and molecular biological levels.

However, it has not always possible using conventional isolation procedures to routinely and reproducibly isolate viral strains from the above mentioned donor or recipients which has frustrated efforts to investigate the cause of the asymptomatic individuals. In accordance with the present invention, methods have now been established to isolate viruses from the above individuals. It has been determined, in accordance with the

- 5 -

present invention, that the six individuals of the cohort are infected by non-pathogenic strains of HIV-1. Furthermore, the non-pathogenic strains of HIV-1 carry one or more nucleotide mutations. The non-pathogenic strains of the present invention enable the generation of a range of therapeutic, diagnostic and targeting agents against HIV-1 infection. The present invention also enables the attenuation of previously pathogenic strains of HIV-1.

Accordingly, one aspect of the present invention contemplates a non-pathogenic isolate of HIV-1 or a component, part, fragment or derivative thereof.

10

In a related embodiment, there is provided a novel isolate of HIV-1 or a component, part, fragment or derivative thereof wherein said HIV-1 isolate is capable of stimulating in a human or primate subject an immune response such as a humoral immune response to at least one HIV-1 glycoprotein such as but not limited to gp41-45, gp120 and/or gp160 while not substantially reducing in said human or primate subject proliferative responses and cytokine production to a mitogen, alloantigen and/or recall antigen compared to a healthy, non-infected human or primate subject. Preferably, the cytokine is IL-2. Preferably, the mitogen is ConA or PHA and the recall antigen is influenza or tetanus toxoid. Preferably, the HIV-1 isolate is non-pathogenic.

20

More particularly, the present invention relates to an isolated HIV-1 strain which:

- (i) is substantially non-pathogenic in human subjects; and
- (ii) carries one or more mutations in its genome resulting in the inability to direct synthesis of at least one pathogenic HIV-1-derived polypeptide or protein.

25

Even more particularly, the present invention provides an isolated HIV-1 strain which:

- (i) is substantially non-pathogenic in human subjects; and
- (ii) carries a mutation in the *nef* gene and/or a long terminal repeat (LTR) region or in a functionally equivalent location in the HIV-1 genome.

30

- 6 -

Still even more particularly, the present invention is directed to an isolated virus which:

- (i) has a genome which is capable of hybridising under medium stringency conditions to complementary nucleic acid from a pathogenic strain of HIV-1;
- 5 (ii) is substantially non-pathogenic in human subjects;
- (iii) carries one or more deletion mutations in a region of its genome corresponding to a *nef* gene in said pathogenic strain of HIV-1; and
- (iv) optionally carries a mutation in one or both LTR regions.

10 In a related embodiment, there is provided an isolated virus which:

- (i) has a genome which is capable of hybridising under medium stringency conditions to complementary nucleic acid from a pathogenic strain of HIV-1;
- (ii) is substantially non-pathogenic in human subjects;
- (iii) carries one or more deletion mutations in an LTR region of its genome; and
- 15 (iv) optionally carries a mutation in a region corresponding to a *nef* gene in said pathogenic strain of HIV-1.

In a further related embodiment, there is provided an isolated virus which:

- (i) has a genome which is capable of hybridising under medium stringency
- 20 conditions to complementary nucleic acid from a pathogenic strain of HIV-1;
- (ii) is substantially non-pathogenic in human subjects; and
- (iii) carries one or more deletion mutations in a region of its genome corresponding to a region which contains *nef* coding sequences and LTR nucleotide sequences.

25 In a particularly preferred embodiment, the present invention provides non-pathogenic HIV-1 isolate C18 deposited at the ECACC on 17 October, 1994 under Provisional Accession Number V94101706.

In a related embodiment, the present invention provides non-pathogenic HIV-1 isolate

30 C98 deposited at the ECACC on 31 October, 1994 under Provisional Accession Number V941031169.

- 7 -

In another embodiment, the present invention provides non-pathogenic HIV-1 isolate C54 deposited at ECACC on 14 February, 1995 under Provisional Accession No. \_\_\_\_\_.

- Although pathogenicity is a relative term, it is used herein in relation to the capacity of a strain of HIV-1 to induce AIDS or AIDS-related disorders in an individual over time. Accordingly, a "non-pathogenic" strain of HIV-1 is a strain which, at the clinical level, does not lead to the development of AIDS, at least within the median time of 6-10 years following infection with HIV-1. At the laboratory level, a non-pathogenic strain of HIV-1 is considered not to alter CD4+ cell counts or  $\beta_2$ -microglobulin concentrations.
- 10 In addition, a non-pathogenic strain of HIV-1 may not alter CD8+ and CD3+ cell counts and would not alter lymphocyte counts. CD4+:CD8+ ratios also remain unchanged relative to normal non-infected individuals. Furthermore, generally, a non-pathogenic strain of HIV-1 does not induce p24 antigenaemia. A non-pathogenic HIV-1 of the present invention is generally still infectious but individuals infected with the virus
- 15 remain free of symptoms for at least 6-10 years after infection.

- A laboratory classified non-pathogenic strain of HIV-1 may be determined at any time after infection. The term "non-pathogenic" is not to be considered as a strain that is never pathogenic under any conditions as this might depend on the host individual, the
- 20 level of immune responsiveness in that individual and the extent or otherwise of other, for example, immune comprising disorders. Accordingly, a "non-pathogenic" HIV-1 isolate of the present invention may also be considered a "low virulent" strain of the virus. A non-pathogenic strain of HIV-1 as contemplated herein may be isolated from an asymptomatic individual or may be derived from a pathogenic strain by mutation.
- 25 Although the present invention is not to be limited to any particular pathogenic strain of HIV-1, for reference purposes, an example of a pathogenic strain is HIV-1 NL4-3 strain as described by Myers *et al* (1992).

- The non-pathogenic nature of the HIV-1 of the present invention is conveniently
- 30 evidenced by the cohort of seven individuals comprising one donor and six recipients which have remained free of symptoms or signs of HIV-1 infection for greater than the median time of 6-10 years. However, the individuals of the cohort are seropositive for

- 8 -

HIV-1 following infection with the virus as determined by Western blot analysis, genetic analysis (e.g. using PCR techniques). A seropositive individual is one showing reactivity to at least one HIV-1 glycoprotein (such as but not limited to gp 41-45, gp160) and at least three other virus-specific bands.

5

In accordance with the present invention, a non-pathogenic HIV-1 isolate is also one of HIV-1 which:

- (i) induces an immune response in a human or primate subject; and
- (ii) does not substantially reduce proliferative responses or cytokine production in response to mitogen, alloantigen and/or recall antigen relative to a healthy, non-infected subject.

10

Preferably, the immune response such as to a glycoprotein, for example gp41-45, gp120 and/or gp160. Preferably, the cytokine monitored is an interleukin, such as IL-2.

15

Preferably, the recall antigen is influenza or tetanus toxoid. A non-pathogenic HIV-1 isolate is also one which:

- (iii) does not substantially alter proliferative responses or cytokine production in response to antigen in allogeneic mononuclear cells.

- 20 The genomes or complementary DNA therefrom of the non-pathogenic HIV-1 isolates of the present invention are capable of hybridising under medium stringency conditions to the corresponding genome or complementary DNA of a pathogenic strain of HIV-1 (e.g. HIV-1 strain NL4-3). The ability to hybridise to a pathogenic strain of HIV-1 applies to a comparison of the entire genome/complementary DNA of a non-pathogenic strain or a fragment which includes genetic material corresponding to a region of the genome 3' of the *nef* gene in a pathogenic strain of HIV-1.

25

For the purposes of reference only, a suitable genomic nucleotide sequence from a pathogenic HIV-1 strain is set forth in SEQ ID NO: 1 from HIV-1 strain NL4-3 (Mason *et al*, 1992):

30

```

1  TGGAAGGGCTAATTTGGTCCCAAAAAGACAAGAGATCCTTGATCTC
51  ATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTACACACCA
101 GCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTTCAAGTTAGT

```

151 CAGTTGAACCAGAGCAAGTAGAAGAGGCCAAATAAGGAGAGAAGAACAGC  
201 TTGTTACACCCTATGAGCCAGCATGGGATGGAGGACCCGGAGGGAGAAGT  
251 ATTAGTGTGGAAGTTTGACAGCCTCCTAGCATTTTCGTACATGGCCCCGAG  
301 AGCTGCATCCGGAGTACTACAAAGACTGCTGACATCGAGCTTTCTACAAG  
5 351 GGACTTTCCGCTGGGGACTTTCCAGGGAGGTGTGGCCTGGGCGGGACTGG  
401 GGAGTGGCGAGCCCTCAGATGCTACATATAAGCAGCTGCTTTTTGCTGT  
451 ACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTA  
501 ACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTCA  
551 AAGTAGTGTGTGCCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTC  
10 601 AGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGG  
651 GACTTGAAAGCGAAAGTAAAGCCAGAGGAGATCTCTCGACGCAGGACTCG  
701 GCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTA  
751 CGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAG  
801 AGCGTCGGTATTAAGCGGGGGAGAATTAGATAAATGGGAAAAAATTCGGT  
15 851 TAAGGCCAGGGGGAAAGAAACAATATAAACTAAACATATAGTATGGGCA  
901 AGCAGGGAGCTAGAACGATTTCGAGTTAATCCTGGCCTTTTAGAGACATC  
951 AGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAG  
1001 GATCAGAAGAACTTAGATCATTATATAATACAATAGCAGTCCTCTATTGT  
1051 GTGCATCAAAGGATAGATGTAAAAGACACCAAGGAAGCCTTAGATAAGAT  
20 1101 AGAGGAAGAGCAAAACAAGTAAGAAAAAGGCACAGCAAGCAGCAGCTG  
1151 ACACAGGAAACAACAGCCAGGTGAGCCAAAATTACCCTATAGTGCAGAAC  
1201 CTCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGC  
1251 ATGGGTAAAAGTAGTAGAAGAGAAGGCTTTTCAGCCCAGAAGTAATACCCA  
1301 TGTTTTTCAGCATTATCAGAAGGAGCCACCCACAAGATTTAAATACCATG  
25 1351 CTAAACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGAC  
1401 CATCAATGAGGAAGCTGCAGAATGGGATAGATTGCATCCAGTGCATGCAG  
1451 GGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCA  
1501 GGAAC TACTAGTACCCTTCAGGAACAAATAGGATGGATGACACATAATCC  
1551 ACCTATCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAA  
30 1601 ATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGACAA  
1651 GGACCAAAGGAACCCCTTTAGAGACTATGTAGACCGATTCTATAAACTCT  
1701 AAGAGCCGAGCAAGCTTCACAAGAGGTAAAAAATTGGATGACAGAAACCT  
1751 TGTTGGTCCAAAATGCGAACCAGATTGTAAGACTATTTTAAAGCATTG  
1801 GGACCAGGAGCGACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGG  
35 1851 GGGACCCGGCCATAAAGCAAGAGTTTTGGCTGAAGCAATGAGCCAAGTAA  
1901 CAAATCCAGCTACCATAATGATACAGAAAGGCAATTTTAGGAACCAAAGA  
1951 AAGACTGTTAAGTGTTCATTGTGGCAAAGAAGGGCACATAGCCAAAAA  
2001 TTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGGAC

2051 ACCAAATGAAAGATTGTACTGAGAGACAGGCTAATTTTTTTAGGGAAGATC  
2101 TGGCCTTCCCACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGA  
2151 GCCAACAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAACAA  
2201 CTCCCTCTCAGAAGCAGGAGCCGATAGACAAGGAACGTATCCTTTAGCT  
5 2251 TCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGG  
2301 GGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATT  
2351 AGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAA  
2401 TTGGAGGTTTTATCAAAGTAGGACAGTATGATCAGATACTCATAGAAATC  
2451 TGCGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAA  
10 2501 CATAATTGGAAGAAATCTGTTGACTCAGATTGGCTGCACTTTAAATTTTC  
2551 CCATTAGTCCTATTGAGACTGTACCAGTAAAATTAAAGCCAGGAATGGAT  
2601 GGCCCAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAGCATT  
2651 AGTAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTG  
2701 GGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAGAC  
15 2751 AGTACTAAATGGAGAAAATTAGTAGATTTTCAGAGAACTTAATAAGAGAAC  
2801 TCAAGATTTCTGGGAAGTTCAATTAGGAATACCACATCCTGCAGGGTTAA  
2851 AACAGAAAAAATCAGTAACAGTACTGGATGTGGGCGATGCATATTTTTCA  
2901 GTTCCCTTAGATAAAGACTTCAGGAAGTATACTGCATTTACCATACCTAG  
2951 TATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCAC  
20 3001 AGGGATGGAAAGGATCACCAGCAATATTCAGTGTAGCATGACAAAATC  
3051 TTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTCATCTATCAATACAT  
3101 GGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAA  
3151 AAATAGAGGAACTGAGACAACATCTGTTGAGGTGGGGATTTACCACACCA  
3201 GACAAAAACATCAGAAAGAACCTCCATTCTTTGGATGGGTTATGAACT  
25 3251 CCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAGGACA  
3301 GCTGGACTGTCAATGACATACAGAAATTAGTGGGAAAATTGAATTGGGCA  
3351 AGTCAGATTTATGCAGGGATTAAAGTAAGGCAATTATGTAACTTCTTAG  
3401 GGGAACCAAAGCACTAACAGAAGTAGTACCACTAACAGAAGAAGCAGAGC  
3451 TAGAACTGGCAGAAAACAGGGAGATTCTAAAAGAACCGGTACATGGAGTG  
30 3501 TATTATGACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCA  
3551 AGGCCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAA  
3601 CAGGAAAATATGCAAGAATGAAGGGTGCCACACTAATGATGTGAAACAA  
3651 TTAACAGAGGCAGTACAAAAAATAGCCACAGAAAGCATAGTAATATGGGG  
3701 AAAGACTCCTAAATTTAAATTACCCATACAAAAGGAAACATGGGAAGCAT  
35 3751 GGTGGACAGAGTATTGGCAAGCCACCTGGATTCTTGAGTGGGAGTTTGTC  
3801 AATACCCCTCCCTTAGTGAAGTTATGGTACCAGTTAGAGAAAGAACCCAT  
3851 AATAGGAGCAGAACTTTCTATGTAGATGGGGCAGCCAATAGGGAAACTA  
3901 AATTAGGAAAAGCAGGATATGTAAGTACAGAGGAAGACAAAAAGTTGTC



3951 CCCCTAACGGACACAACAAATCAGAAGACTGAGTTACAAGCAATTCATCT  
4001 AGCTTTGCAGGATTCGGGATTAGAAGTAAACATAGTGACAGACTCACAAT  
4051 ATGCATTGGGAATCATTCAAGCACAACCAGATAAGAGTGAATCAGAGTTA  
4101 GTCAGTCAAATAATAGAGCAGTTAATAAAAAAGGAAAAAGTCTACCTGGC  
5 4151 ATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATGGGT  
4201 TGGTCAGTGCTGGAATCAGGAAAGTACTATTTTTAGATGGAATAGATAAG  
4251 GCCCAAGAAGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAG  
4301 TGATTTTAACCTACCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTG  
4351 ATAAATGTCAGCTAAAAGGGGAAGCCATGCATGGACAAGTAGACTGTAGC  
10 4401 CCAGGAATATGGCAGCTAGATTGTACACATTTAGAAGGAAAAGTTATCTT  
4451 GGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTAATTCCAG  
4501 CAGAGACAGGGCAAGAAACAGCATACTTCCTCTTAAATTAGCAGGAAGA  
4551 TGGCCAGTAAAAACAGTACATACAGACAATGGCAGCAATTTACCAGTAC  
4601 TACAGTTAAGGCCGCCTGTTGGTGGGCGGGGATCAAGCAGGAATTTGGCA  
15 4651 TTCCCTACAATCCCCAAAGTCAAGGAGTAATAGAATCTATGAATAAAGAA  
4701 TTAAAGAAAATTATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAGAC  
4751 AGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAGAAAAGGGGGGA  
4801 TTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGAC  
4851 ATACAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAATTTTCG  
20 4901 GGTTTATTACAGGGACAGCAGAGATCCAGTTTGGAAAGGACCAGCAAAGC  
4951 TCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATA  
5001 AAAGTAGTGCCAAGAAGAAAAGCAAAGATCATCAGGGATTATGGAAAACA  
5051 GATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAACACA  
5101 TGGAAAAGATTAGTAAAACACCATATGTATATTTCAAGGAAAGCTAAGGA  
25 5151 CTGGTTTTATAGACATCACTATGAAAGTACTAATCCAAAATAAGTTCAG  
5201 AAGTACACATCCCCTAGGGGATGCTAAATTAGTAATAACAACATATTGG  
5251 GGTCTGCATACAGGAGAAAGAGACTGGCATTGTTGGGTCAGGGAGTCTCCAT  
5301 AGAATGGAGGAAAAAGAGATATAGCACACAAGTAGACCCTGACCTAGCAG  
5351 ACCAACTAATTCATCTGCACTATTTTGATTGTTTTTCAGAATCTGCTATA  
30 5401 AGAAATACCATATTAGGACGTATAGTTAGTCCTAGGTGTGAATATCAAGC  
5451 AGGACATAACAAGGTAGGATCTCTACAGTACTTGGCACTAGCAGCATTAA  
5501 TAAAACCAAAACAGATAAAGCCACCTTTGCCTAGTGTTAGGAAACTGACA  
5551 GAGGACAGATGGAACAAGCCCCAGAAGACCAAGGGCCACAGAGGGAGCCA  
5601 TACAATGAATGGACACTAGAGCTTTTAGAGGAACTTAAGAGTGAAGCTGT  
35 5651 TAGACATTTTCCTAGGATATGGCTCCATAACTTAGGACAACATATCTATG  
5701 AAACCTACGGGGATACTTGGGCAGGAGTGAAGCCATAATAAGAATTCTG  
5751 CAACAACTGCTGTTTATCCATTTCAGAATTGGGTGTCGACATAGCAGAAT  
5801 AGGCGTTACTCGACAGAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAGA

- 12 -

5851 CTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAACTGCTTGTACCAA  
5901 TTGCTATTGTAAAAAGTGTGTGCTTTCATTGCCAAGTTTGTTCATGACAA  
5951 AAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGA  
6001 GCTCATCAGAACAGTCAGACTCATCAAGCTTCTCTATCAAAGCAGTAAGT  
5 6051 AGTACATGTAATGCAACCTATAATAGTAGCAATAGTAGCATTAGTAGTAG  
6101 CAATAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATATAGG  
6151 AAAATATTAAGACAAAGAAAAATAGACAGGTAAATTGATAGACTAATAGA  
6201 AAGAGCAGAAGACAGTGGCAATGAGAGTGAAGGAGAAGTATCAGCACTTG  
6251 TGGAGATGGGGGTGGAAATGGGGCACCATGCTCCTTGGGATATTGATGAT  
10 6301 CTGTAGTGCTACAGAAAAATTGTGGGTACAGTCTATTATGGGGTACCTG  
6351 TGTGGAAGGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCA  
6401 TATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCAC  
6451 AGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAAATTTTA  
6501 ACATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGT  
15 6551 TTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTCTGTGT  
6601 TAGTTTAAAGTGCAGTGTATTGAAGAATGATACTAATACCAATAGTAGTA  
6651 GCGGGAGAATGATAATGGAGAAAGGAGAGATAAAAACTGCTCTTTCAAT  
6701 ATCAGCACAAGCATAAGAGATAAGGTGCAGAAAGAATATGCATTCTTTTA  
6751 TAACTTGATATAGTACCAATAGATAATACCAGCTATAGGTTGATAAGTT  
20 6801 GTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCA  
6851 ATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAA  
6901 TAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTAC  
6951 AATGTACACATGGAATCAGGCCAGTAGTATCAACTCAACTGCTGTAAAT  
7001 GGCAGTCTAGCAGAAGAAGATGTAGTAATTAGATCTGCCAATTTACAGA  
25 7051 CAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATT  
7101 GTACAAGACCCAACAATAACAAGAAAAAGTATCCGTATCCAGAGGGGA  
7151 CCAGGGAGAGCATTGTGTTACAATAGGAAAAATAGGAAATATGAGACAAGC  
7201 ACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAG  
7251 CTAGCAAATTAAGAGAACAAATTGGAAATAATAAAACAATAATCTTTAAG  
30 7301 CAATCCTCAGGAGGGGACCCAGAAATTGTAACGCACAGTTTAAATTGTGG  
7351 AGGGGAATTTTTCTACTGTAATCAACACAACCTGTTTAAATAGTACTTGGT  
7401 TTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGAC  
7451 ACAATCACACTCCCATGCAGAATAAAACAATTTATAAACATGTGGCAGGA  
7501 AGTAGGAAAAGCAATGTATGCCCCCTCCCATCAGTGGACAAATTAGATGTT  
35 7551 CATCAAATATTACTGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC  
7601 AATGGGTCCGAGATCTTCAGACCTGGAGGAGGCGATATGAGGGACAATTG  
7651 GAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAG  
7701 TAGCACCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAGAGCA

7751 GTGGGAATAGGAGCTTTGTTCCCTTGGGTTCTTGGGAGCAGCAGGAAGCAC  
7801 TATGGGCTGCACGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGT  
7851 CTGATATAGTGCAGCAGCAGAACAAATTTGCTGAGGGCTATTGAGGCGCAA  
7901 CAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAACAGCTCCAGGCAAG  
5 7951 AATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTT  
8001 GGGGTTGCTCTGGAAACTCATTTCACCACTGCTGTGCCTTGGGAATGCT  
8051 AGTTGGAGTAATAAATCTCTGGAACAGATTTGGAATAACATGACCTGGAT  
8101 GGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAA  
8151 TTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAA  
10 8201 TTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAAACATAACAAATTGGCT  
8251 GTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAA  
8301 GAATAGTTTTTGTCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATAT  
8351 TCACCATTATCGTTTCAGACCCACCTCCCAATCCCGAGGGGACCCGACAG  
8401 GCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCA  
15 8451 TTCGATTAGTGAACGGATCCTTAGCACTTATCTGGGACGATCTGCGGAGC  
8501 CTGTGCCCTCTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAAC  
8551 GAGGATTGTGGAACCTTCTGGGACGCAGGGGGTGGGAAGCCCTCAAATATT  
8601 GGTGGAATCTCCTACAGTATTGGAGTCAGGAATAAAGAATAGTGCTGTT  
8651 AACTTGCTCAATGCCACAGCCATAGCAGTAGCTGAGGGGACAGATAGGGT  
20 8701 TATAGAAGTATTACAAGCAGCTTATAGAGCTATTCGCCACATACCTAGAA  
8751 GAATAAGACAGGGCTTGGAAAGGATTTTGCTATAAGATGGGTGGCAAGTG  
8801 GTCAAAAAGTAGTGTGATTGGATGGCCTGCTGTAAGGGAAAGAATGAGAC  
8851 GAGCTGAGCCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGACCTAGAA  
8901 AAACATGGAGCAATCACAAGTAGCAATACAGCAGCTAACAATGCTGCTTG  
25 8951 TGCCTGGCTAGAAGCACAAGAGGAGGAAGAGGTGGGTTTTCCAGTCACAC  
9001 CTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGC  
9051 CACTTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAAG  
9101 AAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCC  
9151 CTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGACC  
30 9201 TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGA  
9251 GGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATG  
9301 GAATGGATGACCCTGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGC  
9351 CTAGCATTTTCATCACGTGGCCCGAGAGCTGCATCCGAGTACTTCAAGAA  
9401 CTGCTGACATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAG  
35 9451 GGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTGC  
9501 ATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGA  
9551 TCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCT  
9601 CAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTG

- 14 -

9651 TGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAA  
9701 TCTCTAGCA

However, for the purposes of comparing the nucleotide sequences of non-pathogenic  
5 HIV-1 strains including the ability to hybridise to a reference strain, the present  
invention extends to a genomic nucleotide sequence from any pathogenic strain of HIV-  
1.

Accordingly, in a particularly preferred embodiment, there is provided a viral isolate  
10 which:

- (i) carries a genome which is capable of hybridising under medium stringency  
conditions to SEQ ID NO: 1 or a complementary form thereof or an analogous  
sequence from another pathogenic strain of HIV-1; and
- (ii) carries a deletion mutation in a region corresponding to the *nef* gene and/or in  
15 an LTR region. Generally, such an HIV-1 isolate is non-pathogenic as  
hereinbefore defined.

In a related embodiment, there is provided an isolated virus which:

- (i) has a genome which is capable of hybridising under medium stringency  
20 conditions to complementary nucleic acid from a pathogenic strain of HIV-1;  
and
- (ii) carries one or more deletion mutations in a region of its genome corresponding  
to a region which contains *nef* coding sequences and LTR nucleotide sequences.

25 For the purposes of defining the level of stringency, reference can conveniently be made  
to Maniatis *et al* (1982) at pages 387-389 which is herein incorporated by reference  
where the washing steps disclosed are considered high stringency. A low stringency is  
defined herein as being in 1-3X SSC/0.1-0.5% w/v SDS at 37-50°C for 2-3 hours.  
Depending on the source and concentration of nucleic acid involved in the hybridisation,  
30 alternative conditions of stringency may be employed such as medium stringent  
conditions which are considered herein to be 0.1-1X SSC/0.25-0.5% w/v SDS at  $\geq 45^{\circ}\text{C}$   
for 2-3 hours or high stringent conditions considered herein to be 0.1-1X SSC/0.1% w/v  
SDS at  $60^{\circ}\text{C}$  for 1-3 hours.

- 15 -

In a particularly preferred embodiment of the present invention, the non-pathogenic strain of HIV-1 carries a mutation in the *nef* gene and/or LTR region of the genome.

A "mutation" is considered herein to include a single or multiple nucleotide substitution, deletion and/or addition. Most preferred mutations are single or multiple deletions of at least one, most preferably at least ten and even more preferably at least twenty contiguous nucleotides from a region corresponding to the *nef* gene and/or the LTR region. When the non-pathogenic virus carries a mutation in the LTR region, this generally occurs 5' of the Sp1 sites.

10

According to a preferred aspect of the present invention, there is provided a viral isolate which:

- (i) is reactive to antibodies to a glycoprotein from HIV-1 such as at least one of gp41-45, gp120 and/or gp160;
- 15 (ii) is substantially non-pathogenic in human subjects; and
- (iii) carries a deletion mutation of at least ten nucleotides in a region corresponding to the *nef* gene and/or LTR region of a pathogenic strain of HIV-1.

In another embodiment, there is provided a viral isolate which:

- 20 (i) is capable of inducing an immune response to at least one of *gag*, *pol* and/or *env*;
- (ii) is substantially non-pathogenic in human subjects; and
- (iii) carries a deletion mutation of at least ten nucleotides in a region corresponding to the *nef* gene and/or LTR region of a pathogenic strain of HIV-1.

25 Preferably, in respect of the latter embodiment, the immune response is an antibody or a cell mediated response. In a most preferred embodiment, the immune response is a humoral immune response.

The nucleotide sequence of the *nef* gene in HIV-1 NL4-3 is defined in SEQ ID NO:  
30 650:

ATGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCCTGCTGTAAGGGAAAGAAT  
GAGAÇGAGCTGAGCCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGACCTAGAAAAAC  
ATGGAGCAATCACAAGTAGCAATACAGCAGCTAACAATGCTGCTTGTGCCTGGCTAGAA

- 16 -

GCACAAGAGGAGGAAGAGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAAT  
 GACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAG  
 GGCTAATTCACCTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAA  
 GGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTGAGATATCCACTGAC  
 5 CTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATA  
 AAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGAATGGATGACCCTGAG  
 AGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTTCATCACGTGGCCCGAGA  
 GCTGCATCCGGAGTACTTCAAGAACTGCTGA

- 10 The present invention extends to any or all single or multiple nucleotide deletions to a contiguous series of at least ten nucleotides from the *nef* gene which render the strain avirulent. The deletions may encompass the entire gene or parts thereof and may represent a single deletion or two or more deletions. Put in alternative terms, the non-pathogenic HIV-1 isolates of the present invention comprise a nucleotide sequence at the
- 15 corresponding *nef* gene region non-identifiable to SEQ ID NO: 650, said non-identity comprising at least 5%, more preferably at least 10% and even more preferably at least 20% variation thereon.

In a preferred embodiment, therefore, the present invention contemplates a viral isolate  
 20 which:

- (i) is reactive to antibodies to a glycoprotein from HIV-1 such as at least one of gp41-45, gp120 and/or gp160;
- (ii) carries a genome or a part or fragment thereof capable of hybridising under medium stringency conditions to a nucleotide sequence as set forth in SEQ ID  
 25 NO: 1 or a complementary form thereof or an analogous sequence from another pathogenic strain of HIV-1;
- (iii) carries a deletion of at least ten nucleotides in a region corresponding to the *nef* gene in HIV-1 NL4-3; and

wherein said deletion encompasses one or more of the following decanucleotides from  
 30 the *nef* gene of HIV-1 NL4-3 or corresponding sequences from another pathogenic strain of HIV-1:

ATGGGTGGCA (SEQ ID NO: 2);      TGGGTGGCAA (SEQ ID NO: 3);  
 GGGTGGCAAG (SEQ ID NO: 4);      GGTGGCAAGT (SEQ ID NO: 5);  
 GTGGCAAGTG (SEQ ID NO: 6);      TGGCAAGTGG (SEQ ID NO: 7);

- 17 -

GGCAAGTGGT (SEQ ID NO: 8); GCAAGTGGTC (SEQ ID NO: 9);  
 CAAGTGGTCA (SEQ ID NO: 10); AAGTGGTCAA (SEQ ID NO: 11);  
 AGTGGTCAAA (SEQ ID NO: 12); GTGGTCAAAA (SEQ ID NO: 13);  
 TGGTCAAAAA (SEQ ID NO: 14); GGTCAAAAAG (SEQ ID NO: 15);  
 5 GTCAAAAAGT (SEQ ID NO: 16); TCAAAAAGTA (SEQ ID NO: 17);  
 CAAAAAGTAG (SEQ ID NO: 18); AAAAAGTAGT (SEQ ID NO: 19);  
 AAAAGTAGTG (SEQ ID NO: 20); AAAGTAGTGT (SEQ ID NO: 21);  
 AAGTAGTGTG (SEQ ID NO: 22); AGTAGTGTGA (SEQ ID NO: 23);  
 GTAGTGTGAT (SEQ ID NO: 24); TAGTGTGATT (SEQ ID NO: 25);  
 10 AGTGTGATTG (SEQ ID NO: 26); GTGTGATTGG (SEQ ID NO: 27);  
 TGTGATTGGA (SEQ ID NO: 28); GTGATTGGAT (SEQ ID NO: 29);  
 TGATTGGATG (SEQ ID NO: 30); GATTGGATGG (SEQ ID NO: 31);  
 ATTGGATGGC (SEQ ID NO: 32); TTGGATGGCC (SEQ ID NO: 33);  
 TGGATGGCCT (SEQ ID NO: 34); GGATGGCCTG (SEQ ID NO: 35);  
 15 GATGGCCTGC (SEQ ID NO: 36); ATGGCCTGCT (SEQ ID NO: 37);  
 TGGCCTGCTG (SEQ ID NO: 38); GGCCTGCTGT (SEQ ID NO: 39);  
 GCCTGCTGTA (SEQ ID NO: 40); CCTGCTGTAA (SEQ ID NO: 41);  
 CTGCTGTAAG (SEQ ID NO: 42); TGCTGTAAGG (SEQ ID NO: 43);  
 GCTGTAAGGG (SEQ ID NO: 44); CTGTAAGGGA (SEQ ID NO: 45);  
 20 TGTAAGGGAA (SEQ ID NO: 46); GTAAGGGAAA (SEQ ID NO: 47);  
 TAAGGGAAAG (SEQ ID NO: 48); AAGGGAAAGA (SEQ ID NO: 49);  
 AGGGAAAGAA (SEQ ID NO: 50); GGGAAAGAAT (SEQ ID NO: 51);  
 GGAAAGAATG (SEQ ID NO: 52); GAAAGAATGA (SEQ ID NO: 53);  
 AAAGAATGAG (SEQ ID NO: 54); AAGAATGAGA (SEQ ID NO: 55);  
 25 AGAATGAGAC (SEQ ID NO: 56); GAATGAGACG (SEQ ID NO: 57);  
 AATGAGACGA (SEQ ID NO: 58); ATGAGACGAG (SEQ ID NO: 59);  
 TGAGACGAGC (SEQ ID NO: 60); GAGACGAGCT (SEQ ID NO: 61);  
 AGACGAGCTG (SEQ ID NO: 62); GACGAGCTGA (SEQ ID NO: 63);  
 ACGAGCTGAG (SEQ ID NO: 64); CGAGCTGAGC (SEQ ID NO: 65);  
 30 GAGCTGAGCC (SEQ ID NO: 66); AGCTGAGCCA (SEQ ID NO: 67);  
 GCTGAGCCAG (SEQ ID NO: 68); CTGAGCCAGC (SEQ ID NO: 69);  
 TGAGCCAGCA (SEQ ID NO: 70); GAGCCAGCAG (SEQ ID NO: 71);  
 AGCCAGCAGC (SEQ ID NO: 72); GCCAGCAGCA (SEQ ID NO: 73);  
 CCAGCAGCAG (SEQ ID NO: 74); CAGCAGCAGA (SEQ ID NO: 75);  
 35 AGCAGCAGAT (SEQ ID NO: 76); GCAGCAGATG (SEQ ID NO: 77);  
 CAGCAGATGG (SEQ ID NO: 78); AGCAGATGGG (SEQ ID NO: 79);  
 GCAGATGGGG (SEQ ID NO: 80); CAGATGGGGT (SEQ ID NO: 81);  
 AGATGGGGTG (SEQ ID NO: 82); GATGGGGTGG (SEQ ID NO: 83);

- 18 -

ATGGGGTGGG (SEQ ID NO: 84); TGGGGTGGGA (SEQ ID NO: 85);  
GGGGTGGGAG (SEQ ID NO: 86); GGGTGGGAGC (SEQ ID NO: 87);  
GGTGGGAGCA (SEQ ID NO: 88); GTGGGAGCAG (SEQ ID NO: 89);  
TGGGAGCAGT (SEQ ID NO: 90); GGGAGCAGTA (SEQ ID NO: 91);  
5 GGAGCAGTAT (SEQ ID NO: 92); GAGCAGTATC (SEQ ID NO: 93);  
AGCAGTATCT (SEQ ID NO: 94); GCAGTATCTC (SEQ ID NO: 95);  
CAGTATCTCG (SEQ ID NO: 96); AGTATCTCGA (SEQ ID NO: 97);  
GTATCTCGAG (SEQ ID NO: 98); TATCTCGAGA (SEQ ID NO: 99);  
ATCTCGAGAC (SEQ ID NO: 100); TCTCGAGACC (SEQ ID NO: 101);  
10 CTCGAGACCT (SEQ ID NO: 102); TCGAGACCTA (SEQ ID NO: 103);  
CGAGACCTAG (SEQ ID NO: 104); GAGACCTAGA (SEQ ID NO: 105);  
AGACCTAGAA (SEQ ID NO: 106); GACCTAGAAA (SEQ ID NO: 107);  
ACCTAGAAAA (SEQ ID NO: 108); CCTAGAAAAA (SEQ ID NO: 109);  
CTAGAAAAAC (SEQ ID NO: 110); TAGAAAAACA (SEQ ID NO: 111);  
15 AGAAAAACAT (SEQ ID NO: 112); GAAAAACATG (SEQ ID NO: 113);  
AAAAACATGG (SEQ ID NO: 114); AAAACATGGA (SEQ ID NO: 115);  
AAACATGGAG (SEQ ID NO: 116); AACATGGAGC (SEQ ID NO: 117);  
ACATGGAGCA (SEQ ID NO: 118); CATGGAGCAA (SEQ ID NO: 119);  
ATGGAGCAAT (SEQ ID NO: 120); TGGAGCAATC (SEQ ID NO: 121);  
20 GGAGCAATCA (SEQ ID NO: 122); GAGCAATCAC (SEQ ID NO: 123);  
AGCAATCACA (SEQ ID NO: 124); GCAATCACAA (SEQ ID NO: 125);  
CAATCACAAG (SEQ ID NO: 126); AATCACAAGT (SEQ ID NO: 127);  
ATCACAAGTA (SEQ ID NO: 128); TCACAAGTAG (SEQ ID NO: 129);  
CACAAGTAGC (SEQ ID NO: 130); ACAAGTAGCA (SEQ ID NO: 131);  
25 CAAGTAGCAA (SEQ ID NO: 132); AAGTAGCAAT (SEQ ID NO: 133);  
AGTAGCAATA (SEQ ID NO: 134); GTAGCAATAC (SEQ ID NO: 135);  
TAGCAATACA (SEQ ID NO: 136); AGCAATACAG (SEQ ID NO: 137);  
GCAATACAGC (SEQ ID NO: 138); CAATACAGCA (SEQ ID NO: 139);  
AATACAGCAG (SEQ ID NO: 140); ATACAGCAGC (SEQ ID NO: 141);  
30 TACAGCAGCT (SEQ ID NO: 142); ACAGCAGCTA (SEQ ID NO: 143);  
CAGCAGCTAA (SEQ ID NO: 144); AGCAGCTAAC (SEQ ID NO: 145);  
GCAGCTAACA (SEQ ID NO: 146); CAGCTAACAA (SEQ ID NO: 147);  
AGCTAACAAT (SEQ ID NO: 148); GCTAACAATG (SEQ ID NO: 149);  
CTAACAATGC (SEQ ID NO: 150); TAACAATGCT (SEQ ID NO: 151);  
35 AACAATGCTG (SEQ ID NO: 152); ACAATGCTGC (SEQ ID NO: 153);  
CAATGCTGCT (SEQ ID NO: 154); AATGCTGCTT (SEQ ID NO: 155);  
ATGCTGCTTG (SEQ ID NO: 156); TGCTGCTTGT (SEQ ID NO: 157);  
GCTGCTTGTG (SEQ ID NO: 158); CTGCTTGTGC (SEQ ID NO: 159);



- 19 -

TGCTTGTGCC (SEQ ID NO: 160); GCTTGTGCCT (SEQ ID NO: 161);  
CTTGTGCCTG (SEQ ID NO: 162); TTGTGCCTGG (SEQ ID NO: 163);  
TGTGCCTGGC (SEQ ID NO: 164); GTGCCTGGCT (SEQ ID NO: 165);  
TGCCTGGCTA (SEQ ID NO: 166); GCCTGGCTAG (SEQ ID NO: 167);  
5 CCTGGCTAGA (SEQ ID NO: 168); CTGGCTAGAA (SEQ ID NO: 169);  
TGGCTAGAAG (SEQ ID NO: 170); GGCTAGAAGC (SEQ ID NO: 171);  
GCTAGAAGCA (SEQ ID NO: 172); CTAGAAGCAC (SEQ ID NO: 173);  
TAGAAGCACA (SEQ ID NO: 174); AGAAGCACAA (SEQ ID NO: 175);  
GAAGCACAAG (SEQ ID NO: 176); AAGCACAAGA (SEQ ID NO: 177);  
10 AGCACAAGAG (SEQ ID NO: 178); GCACAAGAGG (SEQ ID NO: 179);  
CACAAGAGGA (SEQ ID NO: 180); ACAAGAGGAG (SEQ ID NO: 181);  
CAAGAGGAGG (SEQ ID NO: 182); AAGAGGAGGA (SEQ ID NO: 183);  
AGAGGAGGAA (SEQ ID NO: 184); GAGGAGGAAG (SEQ ID NO: 185);  
AGGAGGAAGA (SEQ ID NO: 186); GGAGGAAGAG (SEQ ID NO: 187);  
15 GAGGAAGAGG (SEQ ID NO: 188); AGGAAGAGGT (SEQ ID NO: 189);  
GGAAGAGGTG (SEQ ID NO: 190); GAAGAGGTGG (SEQ ID NO: 191);  
AAGAGGTGGG (SEQ ID NO: 192); AGAGGTGGGT (SEQ ID NO: 193);  
GAGGTGGGTT (SEQ ID NO: 194); AGGTGGGTTT (SEQ ID NO: 195);  
GGTGGGTTTT (SEQ ID NO: 196); GTGGGTTTTT (SEQ ID NO: 197);  
20 TGGGTTTTTCC (SEQ ID NO: 198); GGGTTTTTCCA (SEQ ID NO: 199);  
GGTTTTTCCAG (SEQ ID NO: 200); GTTTTTCCAGT (SEQ ID NO: 201);  
TTTTTCCAGTC (SEQ ID NO: 202); TTTCCAGTCA (SEQ ID NO: 203);  
TTCCAGTCAC (SEQ ID NO: 204); TCCAGTCACA (SEQ ID NO: 205);  
CCAGTCACAC (SEQ ID NO: 206); CAGTCACACC (SEQ ID NO: 207);  
25 AGTCACACCT (SEQ ID NO: 208); GTCACACCTC (SEQ ID NO: 209);  
TCACACCTCA (SEQ ID NO: 210); CACACCTCAG (SEQ ID NO: 211);  
ACACCTCAGG (SEQ ID NO: 212); CACCTCAGGT (SEQ ID NO: 213);  
ACCTCAGGTA (SEQ ID NO: 214); CCTCAGGTAC (SEQ ID NO: 215);  
CTCAGGTACC (SEQ ID NO: 216); TCAGGTACCT (SEQ ID NO: 217);  
30 CAGGTACCTT (SEQ ID NO: 218); AGGTACCTTT (SEQ ID NO: 219);  
GGTACCTTTA (SEQ ID NO: 220); GTACCTTTAA (SEQ ID NO: 221);  
TACCTTTAAG (SEQ ID NO: 222); ACCTTTAAGA (SEQ ID NO: 223);  
CCTTTAAGAC (SEQ ID NO: 224); CTTTAAGACC (SEQ ID NO: 225);  
TTTAAGACCA (SEQ ID NO: 226); TTAAGACCAA (SEQ ID NO: 227);  
35 TAAGACCAAT (SEQ ID NO: 228); AAGACCAATG (SEQ ID NO: 229);  
AGACCAATGA (SEQ ID NO: 230); GACCAATGAC (SEQ ID NO: 231);  
ACCAATGACT (SEQ ID NO: 232); CCAATGACTT (SEQ ID NO: 233);  
CAATGACTTA (SEQ ID NO: 234); AATGACTTAC (SEQ ID NO: 235);

- 20 -

ATGACTTACA (SEQ ID NO: 236); TGACTTACAA (SEQ ID NO: 237);  
GACTTACAAG (SEQ ID NO: 238); ACTTACAAGG (SEQ ID NO: 239);  
CTTACAAGGC (SEQ ID NO: 240); TTACAAGGCA (SEQ ID NO: 241);  
TACAAGGCAG (SEQ ID NO: 242); ACAAGGCAGC (SEQ ID NO: 243);  
5 CAAGGCAGCT (SEQ ID NO: 244); AAGGCAGCTG (SEQ ID NO: 245);  
AGGCAGCTGT (SEQ ID NO: 246); GGCAGCTGTA (SEQ ID NO: 247);  
GCAGCTGTAG (SEQ ID NO: 248); CAGCTGTAGA (SEQ ID NO: 249);  
AGCTGTAGAT (SEQ ID NO: 250); GCTGTAGATC (SEQ ID NO: 251);  
CTGTAGATCT (SEQ ID NO: 252); TGTAGATCTT (SEQ ID NO: 253);  
10 GTAGATCTTA (SEQ ID NO: 254); TAGATCTTAG (SEQ ID NO: 255);  
AGATCTTAGC (SEQ ID NO: 256); GATCTTAGCC (SEQ ID NO: 257);  
ATCTTAGCCA (SEQ ID NO: 258); TCTTAGCCAC (SEQ ID NO: 259);  
CTTAGCCACT (SEQ ID NO: 260); TTAGCCACTT (SEQ ID NO: 261);  
TAGCCACTTT (SEQ ID NO: 262); AGCCACTTTT (SEQ ID NO: 263);  
15 GCCACTTTTT (SEQ ID NO: 264); CCACTTTTTA (SEQ ID NO: 265);  
CACTTTTTAA (SEQ ID NO: 266); ACTTTTTTAA (SEQ ID NO: 267);  
CTTTTTTAAA (SEQ ID NO: 268); TTTTTTAAAG (SEQ ID NO: 269);  
TTTTTAAAGA (SEQ ID NO: 270); TTTAAAGAAA (SEQ ID NO: 271);  
TTAAAGAAAA (SEQ ID NO: 272); TAAAGAAAAA (SEQ ID NO: 273);  
20 AAAAGAAAAG (SEQ ID NO: 274); AAAGAAAAGG (SEQ ID NO: 275);  
AAGAAAAGGG (SEQ ID NO: 276); AGAAAAGGGG (SEQ ID NO: 277);  
GAAAAGGGGG (SEQ ID NO: 278); AAAAGGGGGG (SEQ ID NO: 279);  
AAAGGGGGGA (SEQ ID NO: 280); AAGGGGGGAC (SEQ ID NO: 281);  
AGGGGGGACT (SEQ ID NO: 282); GGGGGGACTG (SEQ ID NO: 283);  
25 GGGGGACTGG (SEQ ID NO: 284); GGGGACTGGA (SEQ ID NO: 285);  
GGGACTGGAA (SEQ ID NO: 286); GGACTGGAAG (SEQ ID NO: 287);  
GACTGGAAGG (SEQ ID NO: 288); ACTGGAAGGG (SEQ ID NO: 289);  
CTGGAAGGGC (SEQ ID NO: 290); TGAAGGGGCT (SEQ ID NO: 291);  
GGAAGGGCTA (SEQ ID NO: 292); GAAGGGCTAA (SEQ ID NO: 293);  
30 AAGGGCTAAT (SEQ ID NO: 294); AGGGCTAATT (SEQ ID NO: 295);  
GGGCTAATTC (SEQ ID NO: 296); GGCTAATTCA (SEQ ID NO: 297);  
GCTAATTCAC (SEQ ID NO: 298); CTAATTCACT (SEQ ID NO: 299);  
TAATTCACTC (SEQ ID NO: 300); AATTCACTCC (SEQ ID NO: 301);  
ATTCACTCCC (SEQ ID NO: 302); TTCACTCCCA (SEQ ID NO: 303);  
35 TCACTCCCAA (SEQ ID NO: 304); CACTCCCAA (SEQ ID NO: 305);  
ACTCCCAAAG (SEQ ID NO: 306); CTCCCAAAGA (SEQ ID NO: 307);  
TCCCAAAGAA (SEQ ID NO: 308); CCCAAAGAAG (SEQ ID NO: 309);  
CCAAAGAAGA (SEQ ID NO: 310); CAAAGAAGAC (SEQ ID NO: 311);

- 21 -

AAAGAAGACA (SEQ ID NO: 312); AAGAAGACAA (SEQ ID NO: 313);  
AGAAGACAAG (SEQ ID NO: 314); GAAGACAAGA (SEQ ID NO: 315);  
AAGACAAGAT (SEQ ID NO: 316); AGACAAGATA (SEQ ID NO: 317);  
GACAAGATAT (SEQ ID NO: 318); ACAAGATATC (SEQ ID NO: 319);  
5 CAAGATATCC (SEQ ID NO: 320); AAGATATCCT (SEQ ID NO: 321);  
AGATATCCTT (SEQ ID NO: 322); GATATCCTTG (SEQ ID NO: 323);  
ATATCCTTGA (SEQ ID NO: 324); TATCCTTGAT (SEQ ID NO: 325);  
ATCCTTGATC (SEQ ID NO: 326); TCCTTGATCT (SEQ ID NO: 327);  
CCTTGATCTG (SEQ ID NO: 328); CTTGATCTGT (SEQ ID NO: 329);  
10 TTGATCTGTG (SEQ ID NO: 330); TGATCTGTGG (SEQ ID NO: 331);  
GATCTGTGGA (SEQ ID NO: 332); ATCTGTGGAT (SEQ ID NO: 333);  
TCTGTGGATC (SEQ ID NO: 334); CTGTGGATCT (SEQ ID NO: 335);  
TGTGGATCTA (SEQ ID NO: 336); GTGGATCTAC (SEQ ID NO: 337);  
TGGATCTACC (SEQ ID NO: 338); GGATCTACCA (SEQ ID NO: 339);  
15 GATCTACCAC (SEQ ID NO: 340); ATCTACCACA (SEQ ID NO: 341);  
TCTACCACAC (SEQ ID NO: 342); CTACCACACA (SEQ ID NO: 343);  
TACCACACAC (SEQ ID NO: 344); ACCACACACA (SEQ ID NO: 345);  
CCACACACAA (SEQ ID NO: 346); CACACACAAG (SEQ ID NO: 347);  
ACACACAAGG (SEQ ID NO: 348); CACACAAGGC (SEQ ID NO: 349);  
20 ACACAAGGCT (SEQ ID NO: 350); CACAAGGCTA (SEQ ID NO: 351);  
ACAAGGCTAC (SEQ ID NO: 352); CAAGGCTACT (SEQ ID NO: 353);  
AAGGCTACTT (SEQ ID NO: 354); AGGCTACTTC (SEQ ID NO: 355);  
GGCTACTTCC (SEQ ID NO: 356); GCTACTTCCC (SEQ ID NO: 357);  
CTACTTCCCT (SEQ ID NO: 358); TACTTCCCTG (SEQ ID NO: 359);  
25 ACTTCCCTGA (SEQ ID NO: 360); CTTCCCTGAT (SEQ ID NO: 361);  
TTCCCTGATT (SEQ ID NO: 362); TCCCTGATTG (SEQ ID NO: 363);  
CCCTGATTGG (SEQ ID NO: 364); CCTGATTGGC (SEQ ID NO: 365);  
CTGATTGGCA (SEQ ID NO: 366); TGATTGGCAG (SEQ ID NO: 367);  
GATTGGCAGA (SEQ ID NO: 368); ATTGGCAGAA (SEQ ID NO: 369);  
30 TTGGCAGAAC (SEQ ID NO: 370); TGGCAGAACT (SEQ ID NO: 371);  
GGCAGAACTA (SEQ ID NO: 372); GCAGAACTAC (SEQ ID NO: 373);  
CAGAACTACA (SEQ ID NO: 374); AGAACTACAC (SEQ ID NO: 375);  
GAACTACACA (SEQ ID NO: 376); AACTACACAC (SEQ ID NO: 377);  
ACTACACACC (SEQ ID NO: 378); CTACACACCA (SEQ ID NO: 379);  
35 TACACACCAG (SEQ ID NO: 380); ACACACCAGG (SEQ ID NO: 381);  
CACACCAGGG (SEQ ID NO: 382); ACACCAGGGC (SEQ ID NO: 383);  
CACCAGGGCC (SEQ ID NO: 384); ACCAGGGCCA (SEQ ID NO: 385);  
CCAGGGCCAG (SEQ ID NO: 386); CAGGGCCAGG (SEQ ID NO: 387);

- 22 -

AGGGCCAGGG (SEQ ID NO: 388); GGGCCAGGGG (SEQ ID NO: 389);  
GGCCAGGGGT (SEQ ID NO: 390); GCCAGGGGTC (SEQ ID NO: 391);  
CCAGGGGTCA (SEQ ID NO: 392); CAGGGGTCAG (SEQ ID NO: 393);  
AGGGGTCAGA (SEQ ID NO: 394); GGGGTCAGAT (SEQ ID NO: 395);  
5 GGGTCAGATA (SEQ ID NO: 396); GGTCAGATAT (SEQ ID NO: 397);  
GTCAGATATC (SEQ ID NO: 398); TCAGATATCC (SEQ ID NO: 399);  
CAGATATCCA (SEQ ID NO: 400); AGATATCCAC (SEQ ID NO: 401);  
GATATCCACT (SEQ ID NO: 402); ATATCCACTG (SEQ ID NO: 403);  
TATCCACTGA (SEQ ID NO: 404); ATCCACTGAC (SEQ ID NO: 405);  
10 TCCACTGACC (SEQ ID NO: 406); CCACTGACCT (SEQ ID NO: 407);  
CACTGACCTT (SEQ ID NO: 408); ACTGACCTTT (SEQ ID NO: 409);  
CTGACCTTTG (SEQ ID NO: 410); TGACCTTTGG (SEQ ID NO: 411);  
GACCTTTGGA (SEQ ID NO: 412); ACCTTTGGAT (SEQ ID NO: 413);  
CCTTTGGATG (SEQ ID NO: 414); CTTTGGATGG (SEQ ID NO: 415);  
15 TTTGGATGGT (SEQ ID NO: 416); TTGGATGGTG (SEQ ID NO: 417);  
TGGATGGTGC (SEQ ID NO: 418); GGATGGTGCT (SEQ ID NO: 419);  
GATGGTGCTA (SEQ ID NO: 420); ATGGTGCTAC (SEQ ID NO: 421);  
TGGTGCTACA (SEQ ID NO: 422); GGTGCTACAA (SEQ ID NO: 423);  
GTGCTACAAG (SEQ ID NO: 424); TGCTACAAGC (SEQ ID NO: 425);  
20 GCTACAAGCT (SEQ ID NO: 426); CTACAAGCTA (SEQ ID NO: 427);  
TACAAGCTAG (SEQ ID NO: 428); ACAAGCTAGT (SEQ ID NO: 429);  
CAAGCTAGTA (SEQ ID NO: 430); AAGCTAGTAC (SEQ ID NO: 431);  
AGCTAGTACC (SEQ ID NO: 432); GCTAGTACCA (SEQ ID NO: 433);  
CTAGTACCAG (SEQ ID NO: 434); TAGTACCAGT (SEQ ID NO: 435);  
25 AGTACCAGTT (SEQ ID NO: 436); GTACCAGTTG (SEQ ID NO: 437);  
TACCAGTTGA (SEQ ID NO: 438); ACCAGTTGAG (SEQ ID NO: 439);  
CCAGTTGAGC (SEQ ID NO: 440); CAGTTGAGCC (SEQ ID NO: 441);  
AGTTGAGCCA (SEQ ID NO: 442); GTTGAGCCAG (SEQ ID NO: 443);  
TTGAGCCAGA (SEQ ID NO: 444); TGAGCCAGAT (SEQ ID NO: 445);  
30 GAGCCAGATA (SEQ ID NO: 446); AGCCAGATAA (SEQ ID NO: 447);  
GCCAGATAAG (SEQ ID NO: 448); CCAGATAAGG (SEQ ID NO: 449);  
CAGATAAGGT (SEQ ID NO: 450); AGATAAGGTA (SEQ ID NO: 451);  
GATAAGGTAG (SEQ ID NO: 452); ATAAGGTAGA (SEQ ID NO: 453);  
TAAGGTAGAA (SEQ ID NO: 454); AAGGTAGAAG (SEQ ID NO: 455);  
35 AGGTAGAAGA (SEQ ID NO: 456); GGTAAGAAG (SEQ ID NO: 457);  
GTAGAAGAGG (SEQ ID NO: 458); TAGAAGAGGC (SEQ ID NO: 459);  
AGAAGAGGCC (SEQ ID NO: 460); GAAGAGGCCA (SEQ ID NO: 461);  
AAGAGGCCAA (SEQ ID NO: 462); AGAGGCCAAT (SEQ ID NO: 463);

- 23 -

GAGGCCAATA (SEQ ID NO: 464); AGGCCAATAA (SEQ ID NO: 465);  
GGCCAATAAA (SEQ ID NO: 466); GCCAATAAAG (SEQ ID NO: 467);  
CCAATAAAGG (SEQ ID NO: 468); CAATAAAGGA (SEQ ID NO: 469);  
AATAAAGGAG (SEQ ID NO: 470); ATAAAGGAGA (SEQ ID NO: 471);  
5 TAAAGGAGAG (SEQ ID NO: 472); AAAGGAGAGA (SEQ ID NO: 473);  
AAGGAGAGAA (SEQ ID NO: 474); AGGAGAGAAC (SEQ ID NO: 475);  
GGAGAGAACA (SEQ ID NO: 476); GAGAGAACAC (SEQ ID NO: 477);  
AGAGAACACC (SEQ ID NO: 478); GAGAACACCA (SEQ ID NO: 479);  
AGAACACCAG (SEQ ID NO: 480); GAACACCAGC (SEQ ID NO: 481);  
10 AACACCAGCT (SEQ ID NO: 482); ACACCAGCTT (SEQ ID NO: 483);  
CACCAGCTTG (SEQ ID NO: 484); ACCAGCTTGT (SEQ ID NO: 485);  
CCAGCTTGTT (SEQ ID NO: 486); CAGCTTGTTA (SEQ ID NO: 487);  
AGCTTGTTAC (SEQ ID NO: 488); GCTTGTTACA (SEQ ID NO: 489);  
CTTGTTACAC (SEQ ID NO: 490); TTGTTACACC (SEQ ID NO: 491);  
15 TGTTACACCC (SEQ ID NO: 492); GTTACACCCT (SEQ ID NO: 493);  
TTACACCCTG (SEQ ID NO: 494); TACACCCTGT (SEQ ID NO: 495);  
ACACCCTGTG (SEQ ID NO: 496); CACCCTGTGA (SEQ ID NO: 497);  
ACCCTGTGAG (SEQ ID NO: 498); CCCTGTGAGC (SEQ ID NO: 499);  
CCTGTGAGCC (SEQ ID NO: 500); CTGTGAGCCT (SEQ ID NO: 501);  
20 TGTGAGCCTG (SEQ ID NO: 502); GTGAGCCTGC (SEQ ID NO: 503);  
TGAGCCTGCA (SEQ ID NO: 504); GAGCCTGCAT (SEQ ID NO: 505);  
AGCCTGCATG (SEQ ID NO: 506); GCCTGCATGG (SEQ ID NO: 507);  
CCTGCATGGA (SEQ ID NO: 508); CTGCATGGAA (SEQ ID NO: 509);  
TGCATGGAAT (SEQ ID NO: 510); GCATGGAATG (SEQ ID NO: 511);  
25 CATGGAATGG (SEQ ID NO: 512); ATGGAATGGA (SEQ ID NO: 513);  
TGGAATGGAT (SEQ ID NO: 514); GGAATGGATG (SEQ ID NO: 515);  
GAATGGATGA (SEQ ID NO: 516); AATGGATGAC (SEQ ID NO: 517);  
ATGGATGACC (SEQ ID NO: 518); TGGATGACCC (SEQ ID NO: 519);  
GGATGACCCT (SEQ ID NO: 520); GATGACCCTG (SEQ ID NO: 521);  
30 ATGACCCTGA (SEQ ID NO: 522); TGACCCTGAG (SEQ ID NO: 523);  
GACCCTGAGA (SEQ ID NO: 524); ACCCTGAGAG (SEQ ID NO: 525);  
CCCTGAGAGA (SEQ ID NO: 526); CCTGAGAGAG (SEQ ID NO: 527);  
CTGAGAGAGA (SEQ ID NO: 528); TGAGAGAGAA (SEQ ID NO: 529);  
GAGAGAGAAG (SEQ ID NO: 530); AGAGAGAAGT (SEQ ID NO: 531);  
35 GAGAGAAGTG (SEQ ID NO: 532); AGAGAAGTGT (SEQ ID NO: 533);  
GAGAAGTGTT (SEQ ID NO: 534); AGAAGTGTTA (SEQ ID NO: 535);  
GAAGTGTTAG (SEQ ID NO: 536); AAGTGTTAGA (SEQ ID NO: 537);  
AGTGTTAGAG (SEQ ID NO: 538); GTGTTAGAGT (SEQ ID NO: 539);

- 24 -

TGTTAGAGTG (SEQ ID NO: 540); GTTAGAGTGG (SEQ ID NO: 541);  
TTAGAGTGGA (SEQ ID NO: 542); TAGAGTGGAG (SEQ ID NO: 543);  
AGAGTGGAGG (SEQ ID NO: 544); GAGTGGAGGT (SEQ ID NO: 545);  
AGTGGAGGTT (SEQ ID NO: 546); GTGGAGGTTT (SEQ ID NO: 547);  
5 TGGAGGTTTG (SEQ ID NO: 548); GGAGGTTTGA (SEQ ID NO: 549);  
GAGGTTTGAC (SEQ ID NO: 550); AGGTTTGACA (SEQ ID NO: 551);  
GGTTTGACAG (SEQ ID NO: 552); GTTTGACAGC (SEQ ID NO: 553);  
TTTGACAGCC (SEQ ID NO: 554); TTGACAGCCG (SEQ ID NO: 555);  
TGACAGCCGC (SEQ ID NO: 556); GACAGCCGCC (SEQ ID NO: 557);  
10 ACAGCCGCCT (SEQ ID NO: 558); CAGCCGCCTA (SEQ ID NO: 559);  
AGCCGCCTAG (SEQ ID NO: 560); GCCGCCTAGC (SEQ ID NO: 561);  
CCGCCTAGCA (SEQ ID NO: 562); CGCCTAGCAT (SEQ ID NO: 563);  
GCCTAGCATT (SEQ ID NO: 564); CCTAGCATTT (SEQ ID NO: 565);  
CTAGCATTTT (SEQ ID NO: 566); TAGCATTTCA (SEQ ID NO: 567);  
15 AGCATTTTCA (SEQ ID NO: 568); GCATTTTCATC (SEQ ID NO: 569);  
CATTTTCATCA (SEQ ID NO: 570); ATTTTCATCAC (SEQ ID NO: 571);  
TTTCATCACG (SEQ ID NO: 572); TTCATCACGT (SEQ ID NO: 573);  
TCATCACGTG (SEQ ID NO: 574); CATCACGTGG (SEQ ID NO: 575);  
ATCACGTGGC (SEQ ID NO: 576); TCACGTGGCC (SEQ ID NO: 577);  
20 CACGTGGCCC (SEQ ID NO: 578); ACGTGGCCCG (SEQ ID NO: 579);  
CGTGGCCCGA (SEQ ID NO: 580); GTGGCCCGAG (SEQ ID NO: 581);  
TGGCCCGAGA (SEQ ID NO: 582); GGCCCGAGAG (SEQ ID NO: 583);  
GCCCGAGAGC (SEQ ID NO: 584); CCCGAGAGCT (SEQ ID NO: 585);  
CCGAGAGCTG (SEQ ID NO: 586); CGAGAGCTGC (SEQ ID NO: 587);  
25 GAGAGCTGCA (SEQ ID NO: 588); AGAGCTGCAT (SEQ ID NO: 589);  
GAGCTGCATC (SEQ ID NO: 590); AGCTGCATCC (SEQ ID NO: 591);  
GCTGCATCCG (SEQ ID NO: 592); CTGCATCCGG (SEQ ID NO: 593);  
TGCATCCGGA (SEQ ID NO: 594); GCATCCGGAG (SEQ ID NO: 595);  
CATCCGGAGT (SEQ ID NO: 596); ATCCGGAGTA (SEQ ID NO: 597);  
30 TCCGGAGTAC (SEQ ID NO: 598); CCGGAGTACT (SEQ ID NO: 599);  
CGGAGTACTT (SEQ ID NO: 600); GGAGTACTTC (SEQ ID NO: 601);  
GAGTACTTCA (SEQ ID NO: 602); AGTACTTCAA (SEQ ID NO: 603);  
GTACTTCAAG (SEQ ID NO: 604); TACTTCAAGA (SEQ ID NO: 605);  
ACTTCAAGAA (SEQ ID NO: 606); CTTCAAGAAC (SEQ ID NO: 607);  
35 TTCAAGAACT (SEQ ID NO: 608); TCAAGAACTG (SEQ ID NO: 609);  
CAAGAACTGC (SEQ ID NO: 610); AAGAACTGCT (SEQ ID NO: 611);  
AGAACTGCTG (SEQ ID NO: 612); GAACTGCTGA (SEQ ID NO: 613).

Generally, the subject HIV-1 isolate is non-pathogenic as hereinbefore defined. Additionally, reference herein to "a deletion" includes reference to a contiguous or non-contiguous series of two or more deletions.

- 5 The non-pathogenic isolate may carry a single decanucleotide deletion or may carry more than one decanucleotide deletion. Where it carries multiple deletions these may all correspond to a contiguous sequence or be from different parts of the *nef* gene. Furthermore, the terminal end portions of a deletion may lie within a decanucleotide as defined above. It is emphasised that the present invention extends to analogous  
10 sequences from other pathogenic strains of HIV-1 which might carry *nef* genes with a slightly altered sequence relative to HIV-1 strain NL4-3.

In a most preferred embodiment of the present invention, there is provided a non-pathogenic strain of HIV-1 carrying a nucleotide sequence in its genome as set forth  
15 in SEQ ID No. 614:

GAAGAGATTGTTGGGAGAACATGACCTGGATGCAGTGGGAAAAAGAAATTCACAATCACAC  
AAAATACATATACTCCTTACTTGAAAAATCGCAGAACCAACAAGAAAAGAATGAACAAG  
AACTATTGGAATTGGATCAATGGGCAAGTTTGTGGAATTGGTTTGACATAACAAAATGG  
CTGTGGTATATAAAAATATTCATAATGGTAGTAGGAGGCTTGATAGGTTTAAGAATAGT  
20 TTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATACTCACCATTGTCGTTTC  
AGACCCTCCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATCGAAGAAGAAGGT  
GGAGAGAGAGACAGAGACAGATCCACTCGATTAGTACACGGATTCTTAGCACTTTTCTG  
GGACGACCTGAGGAGCCTGTGCCTCTTCCTCTACCACCACTTGAGAGACTTACTCTTGA  
TTGTAACAAGGATTGTGGAACCTCTGGGACGCAGGGGATGGGAAGCCCTCAAATATTGG  
25 TGGAACCTCCTAAAGTATTGGAGCCAGGAACTGCAGAAGAGTGCTGTTATCTTGCTCAA  
TGCCACCGCCATAGCAGTAGCTGAGGGGACAGATAGAGTTTTAGAAGTATTACAAAGAG  
CTTATAGAGCTATCCTCCACATACCTAGAAGAATAAGACAGGGCCTCGAAATGGCTTTG  
CTATAAAATGGGTGGCAAGTGAGCAAAAAGTAGTGTAGTCAGATAGCATGCATCATAAG  
GGGTGGGGGCCAACAATAACAATGCTGATCGTGCCCTGGCTAGAAGCACAAGAGAAGGA  
30 AGAAGCGGGTTTTCCAGTCAAACCTCAGGTAGCTGTAGATCTTAGCCACTTTTTAAAAG  
AAAAGGGGGGACTGGAAGGGCTAATCACTCCCAAAGAAGACAAGATACACAGTGCTGC  
AACTATTACCACTGGAGTCAGCGAAGATAGAAGAGGCCAATGGAGGAGAAAACCACAG  
ATTGTTCTGTTGGGGACTTTCCATCCGTTGGGGACTTTCCAAGGCGGCGTGGCCTGGGT  
GACTAGTTCCGGTGGGGACTTTCCAAGAAGGCGCGCCTGGGCGGGACTGGGGAGTGGC  
35 GAGCCCTCAGATGCTGCATATAAGCAGCTGCTTTCTGCTGTTACTGGGTCTCTCGGGTT

- 26 -

AGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTC  
AATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT  
ATCTAGA;

5 and/or SEQ ID NO: 615:

GAAACAATTTGGGATAACATGACCTGGATGCAGTGGGAAAGAGAAATTGACAATTACAC  
AAACATAATATACACCTTAATTGAAGAATCGCAGAACCAACAAGAAAAAATGAAC TAG  
AATTATTGGAATTGGATAAATGGGCAAATTTGTGGAATTGGTTTAGTATATCAAAC TGG  
CTATGGTATATAAAATTATT CATAATGGTAGTAGGAGGCTTGGTAGGTTTAAAGAATAGT  
10 TTTTACTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATACTCACCATTGTCTGTTTC  
AGACCCACCTCCCAACCCCGAAGGGACCCGACAGGCCAGAAGGAATCGAAGAAGAAGGT  
GGAGAGAGAGACAGAGGCAGCTCCACTCGATTAGTGCACGGATTCTTAGCACTTTTCTG  
GGACGACCTGAGGAGTCTGTGCCTCTTCAGCTACCACCCTTGAGAGACTTACTCTTGA  
TTGTAACGAGGATTGTGGAAC TTCTGGGACGCAGGGGATGGGAAGCCCTCAAATACTGG  
15 TGGAATCTCCTGCAGTATTGGAGGCAGGAAC TACAGAAGAGTGCTGTTAGCTTGTTCAA  
TGGCAGGCCATAGCAGTAGCTGAGGGGACAGATAGAGTTATAGAAGCTTTACGAAGGG  
CTTATAGAGCTATTCTCCACATACCTAGAAGAATAAGACAGGGCTTAGAAAGGGCTTTG  
CTATAAAATGGGTGGCAAGTGGTCAGAAAGTAGTGTGGTTAGAAGGCATGTACCTTTAA  
GACAAGGCAGCTATAGATCTTAGCCGCTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCT  
20 AATTCACCTCACAGAGAAGATCAGTTGAACCAGAAGAAGATAGAAGAGGCCATGAAGAAG  
AAAACAACAGATTGTTCCGTTTGTTCGTTGGGGACTTTCCAGGAGACGTGGCCTGAGT  
GATAAGCCGCTGGGGACTTTCCGAAGAGGCGTGACGGGACTTTCCAAGGCGACGTGGCC  
TGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTGCATATAAGCAGCTGCTTTCTGC  
CTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAG  
25 GGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCC  
CGTCTGTTGTGTGACTCTGGTATCTAGA.

The present invention, however, extends to HIV-1 isolates which are non-pathogenic;  
carry genomes capable of hybridising under low stringency conditions to SEQ ID NO:  
30 614 or SEQ ID NO: 615; and which do not direct the synthesis of a full length *nef* gene  
product.

In a further embodiment the present invention contemplates a viral isolate which:

- (i) is reactive to antibodies to a glycoprotein from HIV-1 such as at least one of  
35 gp41-45, gp120 and/or gp160;
- (ii) carries a genome or a part or fragment thereof capable of hybridising under



- 27 -

medium stringency conditions to a nucleotide sequence as set forth in SEQ ID NO: 1 or a complementary form thereof or an analogous sequence from another pathogenic strain of HIV-1;

(iii) carries a deletion of at least ten nucleotides in a region corresponding to the LTR

5 region in HIV-1 NL4-3; and

wherein said deletion encompasses one or more of the following decanucleotides from the LTR region of HIV-1 NL4-3 or corresponding sequences from another pathogenic strain of HIV-1:

10	GCTTTTGGCC (SEQ ID NO: 652);	CTTTTGGCCT (SEQ ID NO: 653);
	TTTTTGCCTG (SEQ ID NO: 654);	TTTGGCCTGT (SEQ ID NO: 655);
	TTTGCCTGTA (SEQ ID NO: 656);	TTGCCTGTAC (SEQ ID NO: 657);
	TGCCTGTACT (SEQ ID NO: 658);	GCCTGTACTG (SEQ ID NO: 659);
	CCTGTACTGG (SEQ ID NO: 660);	CTGTACTGGG (SEQ ID NO: 661);
15	TGTACTGGGT (SEQ ID NO: 662);	GTACTGGGTC (SEQ ID NO: 663);
	TACTGGGTCT (SEQ ID NO: 664);	ACTGGGTCTC (SEQ ID NO: 665);
	CTGGGTCTCT (SEQ ID NO: 666);	TGGGTCTCTC (SEQ ID NO: 667);
	GGGTCTCTCT (SEQ ID NO: 668);	GGTCTCTCTG (SEQ ID NO: 669);
	GTCTCTCTGG (SEQ ID NO: 670);	TCTCTCTGGT (SEQ ID NO: 671);
20	CTCTCTGGTT (SEQ ID NO: 672);	TCTCTGGTTA (SEQ ID NO: 673);
	CTCTGGTTAG (SEQ ID NO: 674);	TCTCTGGTTA (SEQ ID NO: 675);
	CTGGTTAGAC (SEQ ID NO: 676);	TGGTTAGACC (SEQ ID NO: 677);
	GGTTAGACCA (SEQ ID NO: 678);	GTTAGACCAG (SEQ ID NO: 679);
	TTAGACCAGA (SEQ ID NO: 680);	TAGACCAGAT (SEQ ID NO: 681);
25	AGACCAGATC (SEQ ID NO: 682);	GACCAGATCT (SEQ ID NO: 683);
	ACCAGATCTG (SEQ ID NO: 684);	CCAGATCTGA (SEQ ID NO: 685);
	CAGATCTGAG (SEQ ID NO: 686);	AGATCTGAGC (SEQ ID NO: 687);
	GATCTGAGCC (SEQ ID NO: 688);	ATCTGAGCCT (SEQ ID NO: 689);
	TCTGAGCCTG (SEQ ID NO: 690);	CTGAGCCTGG (SEQ ID NO: 691);
30	TGAGCCTGGG (SEQ ID NO: 692);	GAGCCTGGGA (SEQ ID NO: 693);
	AGCCTGGGAG (SEQ ID NO: 694);	GCCTGGGAGC (SEQ ID NO: 695);
	CCTGGGAGCT (SEQ ID NO: 696);	CTGGGAGCTC (SEQ ID NO: 697);
	TGGGAGCTCT (SEQ ID NO: 698);	GGGAGCTCTC (SEQ ID NO: 699);
	GGAGCTCTCT (SEQ ID NO: 700);	GAGCTCTCTG (SEQ ID NO: 701);
35	AGCTCTCTGG (SEQ ID NO: 702);	GCTCTCTGGC (SEQ ID NO: 703);
	CTCTCTGGCT (SEQ ID NO: 704);	TCTCTGGCTA (SEQ ID NO: 705);

CTCTGGCTAA (SEQ ID NO: 706); TCTGGCTAAC (SEQ ID NO: 707);  
CTGGCTAACT (SEQ ID NO: 708); TGGCTAACTA (SEQ ID NO: 709);  
GGCTAACTAG (SEQ ID NO: 710); GCTAACTAGG (SEQ ID NO: 711);  
CTAACTAGGG (SEQ ID NO: 712); TAACTAGGGA (SEQ ID NO: 713);  
5 AACTAGGGAA (SEQ ID NO: 714); ACTAGGGAAC (SEQ ID NO: 715);  
CTAGGGAACC (SEQ ID NO: 716); TAGGGAACCC (SEQ ID NO: 717);  
AGGGAACCCA (SEQ ID NO: 718); GGAACCCAC (SEQ ID NO: 719);  
GGAACCCACT (SEQ ID NO: 720); GAACCCACTG (SEQ ID NO: 721);  
AACCCACTGC (SEQ ID NO: 722); ACCCACTGCT (SEQ ID NO: 723);  
10 CCCACTGCTT (SEQ ID NO: 724); CCACTGCTTA (SEQ ID NO: 725);  
CACTGCTTAA (SEQ ID NO: 726); ACTGCTTAAG (SEQ ID NO: 727);  
CTGCTTAAGC (SEQ ID NO: 728); TGCTTAAGCC (SEQ ID NO: 729);  
GCTTAAGCCT (SEQ ID NO: 730); CTTAAGCCTC (SEQ ID NO: 731);  
TTAAGCCTCA (SEQ ID NO: 732); TAAGCCTCAA (SEQ ID NO: 733);  
15 AAGCCTCAAT (SEQ ID NO: 734); AGCCTCAATA (SEQ ID NO: 735);  
GCCTCAATAA (SEQ ID NO: 736); CCTCAATAAA (SEQ ID NO: 737);  
CTCAATAAAG (SEQ ID NO: 738); TCAATAAAGC (SEQ ID NO: 739);  
CAATAAAGCT (SEQ ID NO: 740); AATAAAGCTT (SEQ ID NO: 741);  
ATAAAGCTTG (SEQ ID NO: 742); TAAAGCTTGC (SEQ ID NO: 743);  
20 AAAGCTTGCC (SEQ ID NO: 744); AAGCTTGCCCT (SEQ ID NO: 745);  
AGCTTGCCCTT (SEQ ID NO: 746); GCTTGCCCTG (SEQ ID NO: 747);  
CTTGCCCTGA (SEQ ID NO: 748); TTGCCTTGAG (SEQ ID NO: 749);  
TGCCTTGAGT (SEQ ID NO: 750); GCCTTGAGTG (SEQ ID NO: 751);  
CCTTGAGTGC (SEQ ID NO: 752); CTTGAGTGCT (SEQ ID NO: 753);  
25 TTGAGTGCTT (SEQ ID NO: 754); TGAGTGCTTC (SEQ ID NO: 755);  
GAGTGCTTCA (SEQ ID NO: 756); AGTGCTTCAA (SEQ ID NO: 757);  
GTGCTTCAAG (SEQ ID NO: 758); TGCTTCAAGT (SEQ ID NO: 759);  
GCTTCAAGTA (SEQ ID NO: 760); CTTCAAGTAG (SEQ ID NO: 761);  
TTCAAGTAGT (SEQ ID NO: 762); TCAAGTAGTG (SEQ ID NO: 763);  
30 CAAGTAGTGT (SEQ ID NO: 764); AAGTAGTGTG (SEQ ID NO: 765);  
AGTAGTGTGT (SEQ ID NO: 766); GTAGTGTGTG (SEQ ID NO: 767);  
TAGTGTGTGC (SEQ ID NO: 768); AGTGTGTGCC (SEQ ID NO: 769);  
GTGTGTGCCC (SEQ ID NO: 770); TGTGTGCCCCG (SEQ ID NO: 771);  
GTGTGCCCCGT (SEQ ID NO: 772); TGTGCCCCGTC (SEQ ID NO: 773);  
35 GTGCCCCGTCT (SEQ ID NO: 774); TGCCCCGTCTG (SEQ ID NO: 775);  
GCCCGTCTGT (SEQ ID NO: 776); CCCGTCTGTT (SEQ ID NO: 777);  
CCGTCTGTTG (SEQ ID NO: 778); CGTCTGTTGT (SEQ ID NO: 779);  
GTCTGTTGTG (SEQ ID NO: 780); TCTGTTGTGT (SEQ ID NO: 781);

- 29 -

CTGTTGTGTG (SEQ ID NO: 782); TGTGTGTGA (SEQ ID NO: 783);  
 GTTGTGTGAC (SEQ ID NO: 784); TTGTGTGACT (SEQ ID NO: 785);  
 TGTGTGACTC (SEQ ID NO: 786); GTGTGACTCT (SEQ ID NO: 787);  
 TGTGTGACTC (SEQ ID NO: 788); GTGTGACTCT (SEQ ID NO: 789);  
 5 TGTGACTCTG (SEQ ID NO: 790); GTGACTCTGG (SEQ ID NO: 791);  
 TGACTCTGGT (SEQ ID NO: 792); GACTCTGGTA (SEQ ID NO: 793);  
 ACTCTGGTAA (SEQ ID NO: 794); CTCTGGTAAC (SEQ ID NO: 795);  
 TCTGGTAACT (SEQ ID NO: 796); CTGGTAACTA (SEQ ID NO: 797);  
 TGGTAACTAG (SEQ ID NO: 798); GGTAACTAGA (SEQ ID NO: 799).

10

The non-pathogenic isolate may carry a single decanucleotide deletion in the LTR region or may carry multiple deletions in the same region or in the LTR region and another region such as the *nef* gene. In particular, the mutation may be in the LTR/*nef* overlap region. Where it carries multiple deletions, these may correspond to a contiguous  
 15 sequence or be from different parts of the LTR region and/or *nef* gene. Furthermore, the terminal end portions of a deletion may lie within a decanucleotide as defined above.

Yet another aspect of the present invention provides an infectious molecular clone comprising genetic sequences derived from the non-pathogenic HIV-1 isolates as  
 20 hereinbefore described and includes genetic sequences encoding major structural proteins such as *gag*, *env* and *pol*. Infectious molecular clones are particularly useful as genetic compositions capable of "infecting" host cells without need of viral coat. The infectious molecular clones of the present invention may also be derived from pathogenic HIV-1 strains rendered non-pathogenic as hereindescribed.

25

According to this latter embodiment, there is contemplated a method of attenuating a pathogenic strain of HIV-1, said method comprising inducing a mutation in the *nef* gene and/or an LTR region to generate a non-pathogenic HIV-1 strain as hereinbefore described. Preferred mutations are deletions of at least ten nucleotides such as one or  
 30 more of the decanucleotides as hereinbefore described. The mutations may also constitute substitutions and/or insertions of heterologous nucleic acid molecules in the *nef* and/or LTR regions such as the incorporation of a sense (i.e. co-suppression) antisense or ribozyme molecule (see below).

- 30 -

In still yet another embodiment of the present invention, there is provided an isolated, non-pathogenic strain of HIV-1 comprising a deletion in its genome of at least 10 nucleotide within the region of nucleotide 8787 and 9709 and more particularly within the region 8800 and 9700 and even more preferably within the region 8800 and 9410,  
5 using the nucleotide numbering of HIV-1 NL4-3.

In one embodiment, the deletion of at least 10 nucleotides is from within a region selected from the list consisting of (using the nucleotide numbers of HIV-1 NL4-3):

- |            |       |                |
|------------|-------|----------------|
| nucleotide | (i)   | 8830-8862;     |
| 10         | (ii)  | 9009-9035;     |
|            | (iii) | 9019-9029; and |
|            | (iv)  | 9033-9049.     |

In another embodiment, the deletion of at least 10 nucleotides is from within a region selected from the list consisting of (using the nucleotide numbers of HIV-1 NL4-3):

- |            |        |                |
|------------|--------|----------------|
| nucleotide | (v)    | 9281-9371;     |
|            | (vi)   | 9281-9362;     |
|            | (vii)  | 9105-9224; and |
|            | (viii) | 9271-9370.     |

20

In yet another embodiment, the deletion of at least 10 nucleotides is from within a region selected from the list consisting of (using the nucleotide numbers of HIV-1 NL4-3):

- |            |       |                |
|------------|-------|----------------|
| nucleotide | (ix)  | 8882-8928;     |
| 25         | (x)   | 8850-9006;     |
|            | (xi)  | 8792-9041; and |
|            | (xii) | 9112-9204.     |

In still yet another embodiment, the deletion of at least 10 nucleotides is from within a region selected from the list consisting of (using the nucleotide numbers of HIV-1 NL4-3):

30

- nucleotide (xiii) 9105-9224;  
(xiv) 9389-9395; and  
(xv) 9281-9366.

5

The above embodiments including any combinations thereof or functional equivalents thereof. Most preferred variants of HIV-1 are defined in Table 3.

Reference herein to the non-pathogenic HIV-1 strains of the present invention includes  
10 reference to components, parts, fragments and derivatives thereof including both genetic  
and non-genetic material. Furthermore, the non-pathogenic HIV-1 strains may be in  
isolated form or resident in peripheral blood mononuclear cells (PBMCs) or other like  
cells where the genome of the HIV-1 strains is integrated as DNA from said HIV-1  
strains such as proviral DNA. In addition, the present invention extends to recombinant  
15 virus such as from (or resident in) prokaryotes or eukaryotes as well as in the form of  
infectious molecular clones.

Accordingly, the present invention provides for the non-pathogenic HIV-1 isolate,  
genomic material therefrom, complementary proviral DNA, molecular infectious clones,  
20 recombinant viral particles or genetic sequences therefrom or cells expressing same or  
blood cells carrying proviral DNA or to any mutants, derivatives, components,  
fragments, parts, homologues or analogues of the foregoing.

The non-pathogenic HIV-1 strains of the present invention are particularly useful in the  
25 development of therapeutic compositions, therapeutic molecules and/or diagnostic  
reagents. With regards to the former, the non-pathogenic HIV-1 strain may be  
considered as a live attenuated vaccine where individuals carrying DNA derived from  
said non-pathogenic HIV-1 strain such as proviral DNA in target cells are protected from  
infection by a corresponding pathogenic strain. The term "vaccine" is used in its broadest  
30 sense as a therapeutic composition or molecule which prevents or reduces HIV-1  
infection or risk of infection or which ameliorates the symptoms of infection. It may  
involve the stimulation of an immune response or may involve blocking HIV-1 cells

receptors and/or the use of genetic compositions, for example, to introduce ribozymes or antisense molecules to HIV-1 directed genetic sequences or to prepare infectious molecular clones. For convenience, all such compositions are referred hereinafter to "therapeutic compositions".

5

Accordingly, the present invention contemplates a method for inhibiting or reducing the risk of infection by a pathogenic strain of HIV-1, said method comprising administering to a subject a non-pathogenic HIV-1 as hereinbefore defined in an amount effective to infect target cells and to generate target cells carrying proviral DNA from said non-pathogenic HIV-1. More particularly, the present invention contemplates a method for inhibiting or reducing productive infection of an individual by a pathogenic strain of HIV-1, said method comprising administering to a subject a non-pathogenic isolate of HIV-1 in an amount effective to infect target cells and to generate target cells carrying proviral DNA from said non-pathogenic HIV-1. By "productive infection" as used in the specification and claims herein is meant the infection of a cell or cells by a pathogenic strain of HIV-1 which leads ultimately to the symptomology of AIDS or AIDS related diseases. A cell infected productively produces pathogenic virions. By definition, infection of an individual by a non-pathogenic strain of HIV-1 would not lead to productive infection. Non pathogenic HIV-1 strains generally replicate to a sufficient extent to protect cells against challenge by virulent or pathogenic strains. The methods of the present invention are applicable prophylactically (i.e. to prevent *de novo* infection) or therapeutically (i.e. to reduce or slow disease progression).

The present invention further provides a method for vaccinating an individual against the development of AIDS or AIDS related diseases, said method comprising administering to said individual a non-pathogenic isolate of HIV-1 in an amount effective to infect target cells and to generate target cells carrying proviral DNA from said non-pathogenic HIV-1. The term "vaccinating" should not be taken as limiting the invention to the prevention of HIV-1 infection by solely immunological means. The term "vaccinating" includes any means of preventing productive infection of an individual by pathogenic HIV-1.

As an alternative to the above methods, a therapeutic composition as hereinbefore defined is administered. The non-pathogenic isolate may be administered *inter alia* as an isolated viral preparation or *via* infected blood cells. Another aspect of the invention provides a therapeutic composition for inhibiting or reducing the risk of infection by a pathogenic strain of HIV-1 said therapeutic composition comprising a non-pathogenic strain of HIV-1 or genetic sequences derived therefrom as hereinbefore described and optionally one or more pharmaceutically acceptable carriers and/or diluents. The therapeutic composition of the present invention is generally suitable for intravenous, intraperitoneal, intramuscular, intramucosal (e.g. nasal spray, respiratory spray) or other forms of parenteral administration. The therapeutic composition might also be administered *via* an implant or rectally or orally. In addition to the mutations contemplated above, the non-pathogenic HIV-1 strain may also contain one or more other mutations to further reduce the risk of reversion to virulence and/or to insert a genetic sequence capable of providing directly or indirectly an identifiable signal, having further anti-HIV-1 properties and/or immunostimulatory or cell regulatory properties.

For example, the non-pathogenic HIV-1 isolate in the therapeutic composition may comprise additional genetic material capable of directing expression of antisense nucleotide sequences to inhibit expression of one or more proteins encoded by a pathogenic strain of HIV-1. Alternatively, sense co-suppression may be employed. Preferred sense or antisense molecules would reduce expression of the *nef* gene or affect normal functioning of the LTR region.

According to this embodiment, the non-pathogenic HIV-1 strain may be considered as a targeting agent to introduce genetic constructs capable of reducing expression of one or more HIV-1 proteins or polypeptides. In this embodiment there is provided a viral isolate which:

- (i) is genetically or immunologically related to a pathogenic strain of HIV-1;
- (ii) is substantially non-pathogenic in human subjects;
- (iii) comprises a first nucleotide sequence constituting its genome which is capable of hybridising under medium stringency conditions to SEQ ID NO: 1 or a complementary form thereof; and

- 34 -

- (iv) comprises a second nucleotide sequence within said first nucleotide sequence and which second nucleotide sequence directs expression of a mRNA molecule capable of inhibiting, reducing or otherwise down-regulating translation of a protein or polypeptide encoded by a pathogenic strain of HIV-1 or inhibit, reduce  
5 or otherwise down regulate operation of a non-protein encoding a region of a pathogenic strain of HIV-1.

Preferred proteins or polypeptides targeted for reduced expression are those encoded by one or more of the following: *gag*, *pol*, *env*, *tat*, *rev*, *vpu*, *vpr*, *vif* and/or *nef* genes. The  
10 most preferred protein or polypeptide targeted for reduced expression is the product of the *nef* gene. Alternatively, the target nucleotide sequence does not encode a polypeptide or protein but is required for other functions such as integration or excision from the human genome or expression of genes on the viral genome. An example of such a nucleotide sequence is the LTR region. Accordingly, the present invention  
15 extends to disruption to the function of such regions.

The above aspect relates to use of antisense technology. The present invention extends, however, to use of ribozymes and/or co-suppression to achieve the same results. In an alternative embodiment, or in addition to, the second (or optionally a third) nucleotide  
20 sequence encodes an antiviral agent (e.g. interferon) and/or an immune enhancing agent.

The identification of deletions *inter alia* in the *nef* gene and/or LTR region in asymptomatic subjects provides a unique opportunity to study the *in vivo* effects of attenuated HIV-1 strains carrying one or more mutations in selected genetic regions. In  
25 particular, the present invention provides a means for designing therapeutic compositions directed to inhibiting expression of a *nef* gene and/or LTR region in a pathogenic HIV-1 strain (such as contemplated above) as well as developing a therapeutic regimen aimed at inhibiting the activity of the *nef* gene product for the function of the LTR region. According to this latter embodiment, the present invention provides a therapeutic  
30 composition comprising a molecule capable of inhibiting the intracellular activity of the *nef* gene product and/or LTR region, said composition further comprising one or more pharmaceutically acceptable carriers and/or diluents.



- 35 -

The molecule contemplated by the above aspect of the subject invention may be a protein, polypeptide, peptide, chemical compound, sugar moiety or derivative of the *nef* gene product. The molecule will need to be capable of entering an infected cell. Where the molecule is a protein, polypeptide or peptide, it may be encoded by the second  
5 nucleotide sequence on the targeting vector as contemplated above. Alternatively, the molecule may be a nucleic acid molecule capable of targeting the *nef* gene or LTR region.

The deletion mutants of the present invention may result in a modified *nef* gene product  
10 either having no readily discernable activity or having activity different to the naturally occurring *nef* protein. In any event, if a mutant *nef* gene product is produced, it will generally have a lower molecular weight than the naturally occurring *nef* protein and will have a different overall amino acid sequence. This provides, therefore, for a means for diagnosing individuals with benign HIV-1 infection by, for example, assaying for a  
15 modified *nef* protein or screening for a modified *nef* gene sequence. Alternatively, benign HIV-1 infection may be detected by assaying for a modified LTR region such as an altered nucleotide sequence.

According to one embodiment, there is provided a method for determining the  
20 pathogenicity of an HIV-1 strain after said HIV-1 strain infects cells of an individual, said method comprising contacting a biological sample from said individual with an effective amount of an antibody specific to a *nef* protein from a non-pathogenic strain of HIV-1 (as hereinbefore defined) for a time and under conditions sufficient to form an antibody-modified *nef* protein complex and then detecting said complex. The  
25 presence of said complex is indicative of a modified *nef* gene product and of the non-pathogenicity of the strain of HIV-1. The biological sample is a sample likely to contain the modified *nef* gene product such as tissue extract or cell extract of an infected cell. However, where the modified *nef* gene product is capable of permeation or transport out of the cell, suitable biological fluid would include serum, whole blood, lymph and  
30 mucosal secretion amongst other fluids. Many variations in the subject assay are possible and are contemplated herein. For example, an assay could be based on the inability for a *nef* specific antibody to bind to a modified *nef* protein. For the purposes

of the present invention the term "contacting" including "mixing".

The presence of a modified *nef* molecule in biological fluid can be detected using a wide range of immunoassay techniques such as those described in US Patent Nos. 4,016,043, 4,424,279 and 4,018,653. These include both single-site and two-site, or "sandwich", assays of the non-competitive types, as well as in the traditional competitive binding assays and include ELISA and RIA techniques. Sandwich assays are among the most useful and commonly used assays and are favoured for use in the present invention. A number of variations of the sandwich assay technique exist, and all are intended to be encompassed by the present invention. Briefly, and by way of example only, in a typical forward assay, a modified *nef* product-specific antibody is immobilised onto a solid substrate to form a first complex and the sample to be tested for modified *nef* product brought into contact with the bound molecule. After a suitable period of incubation, for a period of time sufficient to allow formation of an antibody-modified *nef* product secondary complex, a second modified *nef* protein antibody, labelled with a reporter molecule capable of producing a detectable signal, is then added and incubated, allowing time sufficient for the formation of a tertiary complex of antibody-modified *nef* product-antibody. Any unreacted material is washed away, and the presence of the tertiary complex is determined by observation of a signal produced by the reporter molecule. The results may either be qualitative, by simple observation of the visible signal or may be quantitated by comparison with a control sample containing known amounts of hapten. Variations of the forward assay include a simultaneous assay, in which both sample and labelled antibody are added simultaneously to the bound antibody, or a reverse assay in which the labelled antibody and sample to be tested are first combined, incubated and then added simultaneously to the bound antibody. These techniques are well known to those skilled in the art, and the possibility of minor variations will be readily apparent. The antibodies used above may be monoclonal or polyclonal.

The solid substrate is typically glass or a polymer, the most commonly used polymers being cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene. The solid supports may be in the form of tubes, beads, discs or microplates, or any other

- 37 -

surface suitable for conducting an immunoassay. The binding processes are well-known in the art and generally consist of cross-linking covalently binding or physically adsorbing the molecule to the insoluble carrier.

- 5 By "reporter molecule", as used in the present specification, is meant a molecule which, by its chemical nature, produces an analytically identifiable signal which allows the detection of antigen-bound antibody. Detection may be either qualitative or quantitative. The most commonly used reporter molecule in this type of assay are either enzymes, fluorophores or radionuclide containing molecules (i.e. radioisotopes). In the case of an  
10 enzyme immunoassay, an enzyme is conjugated to the second antibody, generally by means of glutaraldehyde or periodate. As will be readily recognised, however, a wide variety of different conjugation techniques exist which are readily available to one skilled in the art. Commonly used enzymes include horseradish peroxidase, glucose oxidase,  $\beta$ -galactosidase and alkaline phosphatase, amongst others. The substrates to be  
15 used with the specific enzymes are generally chosen for the production, upon hydrolysis by the corresponding enzyme, of a detectable colour change. It is also possible to employ fluorogenic substrates, which yield a fluorescent product.

- Alternatively, fluorescent compounds, such as fluorescein and rhodamine, may be  
20 chemically coupled to antibodies without altering their binding capacity. When activated by illumination with light of a particular wavelength, the fluorochrome-labelled antibody adsorbs the light energy, inducing a state of excitability in the molecule, followed by emission of the light at a characteristic colour visually detectable with a light microscope. As in the EIA, the fluorescent labelled antibody is allowed to bind to the  
25 first antibody-hapten complex. After washing off the unbound reagent, the remaining complex is then exposed to the light of the appropriate wavelength, the fluorescence observed indicates the presence of the hapten of interest. Immunofluorescence and EIA techniques are both very well established in the art and are particularly preferred for the present method. However, other reporter molecules, such as radioisotope,  
30 chemiluminescent or bioluminescent molecules, may also be employed. It will be readily apparent to the skilled technician how to vary the procedure to suit the required purpose. It will also be apparent that the foregoing can be used to label a modified *nef*

- 38 -

product and to use same directly in the detection of, for example, circulatory antibodies specific to said modified *nef* product.

Alternatively, genetic assays may be conducted to screen for aberrations in the *nef* gene and/or LTR region. Such a genetic assay may be by Southern or Northern blot analysis, PCR analysis or the like using oligonucleotides specific to a deleted region of a *nef* gene and/or LTR region.

According to this embodiment there is provided a method for determining the pathogenicity of an HIV-1 strain after said HIV-1 strain infects cells of an individual, said method comprising determining directly or indirectly the presence of a deletion mutation in the genome of said HIV-1 wherein the presence of a such a mutation is indicative of the presence of a non-pathogenic strain of HIV-1. The deletion mutation may result in the genome being unable to synthesise a polypeptide or protein from a pathogenic strain of HIV-1 or may direct the synthesis of a truncated form of said polypeptide or protein. The mutation may also lead to altered expression of a polypeptide detectable by, for example, decreased synthesis of a particular protein, such as the *nef* gene product. Alternatively, the deletion mutation affects the LTR region or a regulatory region of the HIV-1 genome. In either case, affected viruses may also be detected by, for example, observing low viral copy numbers such as low viral loads.

Preferably, said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides within the region from nucleotide 8787 to nucleotide 9709 using the nucleotide numbering of HIV-1 strain NL4-3.

25

Preferably, said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

- |            |       |                |
|------------|-------|----------------|
| nucleotide | (i)   | 8830-8862;     |
|            | (ii)  | 9009-9035;     |
| 30         | (iii) | 9019-9029; and |
|            | (iv)  | 9033-9049.     |

- 39 -

Preferably, said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

- 5 nucleotide (v) 9281-9371;  
(vi) 9281-9362;  
(vii) 9105-9224; and  
(viii) 9271-9370.

Preferably, said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

- 10 nucleotide (ix) 8882-8928;  
(x) 8850-9006;  
(xi) 8792-9041; and  
(xii) 9112-9204.

15 Preferably, said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

- nucleotide (xiii) 9105-9224;  
(xiv) 9389-9395; and  
(xv) 9281-9366.

20

The above nucleotide numbers are based on the nucleotide numbering in the NL4-3 genome.

25 Particularly preferred oligonucleotides are based on the deleted regions of the *nef* gene and/or LTR region such as but not limited to one or more oligonucleotides based on SEQ ID NO: 2 to SEQ ID NO: 613 and/or SEQ ID NO: 652 to SEQ ID NO: 799.

30 The present invention further extends to kits for the diagnosis of infection by pathogenic strains of HIV-1 or for determining the pathogenicity of infecting virus. The kits would be in compartmental form each comprising one or more suitable reagents for conducting the assay.

- 40 -

The present invention is further described by the following non-limiting Figures and Examples.

**In the Figures:**

5

Figure 1 is a representation showing the alignment of the nucleotide sequences from donor D36 peripheral blood mononuclear cell (PBMC) [D36P] and non-pathogenic HIV-1 from recipient C18 HIV<sub>StV</sub> [C18S], C18 HIV<sub>MBC</sub> [C18M] and C98 HIV [C98H] and C54 PBMC [C54P] with the equivalent region of HIV-1 NL4-3. Sequences labelled  
10 PBMC are from patient PBMC, those labelled HIV are from virus isolated from patient PBMC and grown in culture. Numbering for HIV-1 NL4-3 is as per Myer *et al* (1992) where nucleotide 1 is the first nucleotide of the complete proviral DNA sequence. D36P, C18S, C18M, C98H and C54P are numbered from the start of the region sequenced. Identity with NL4-3 sequence is shown by (\*). Deleted nucleotides are  
15 shown by (-). Spaces introduced to maximise alignment are shown by (.). Features in NL4-3 are marked by overlining the sequence, features in D36 and C18 are marked by underlining the sequence.

Figure 2 shows the alignment of encoded amino acid sequences of (a) *tat* exon 3 and  
20 (b) *rev* exon 3 from HIV-1 NL4-3, D36 PBMC, C18 HIV<sub>StV</sub> and C98 HIV. In-phase termination codons (\*) and NL4-3 encoded amino acid numbers are shown.

Figure 3 is a representation showing the alignment of C-terminal envelope glycoprotein gp41 amino acid sequences encoded by D36 PMBC, C18 HIV<sub>StV</sub>, C18 HIV<sub>MBC</sub> and  
25 C98 HIV. Numbering is that of the amino acid sequence of the mature envelope glycoprotein of HIV NL4-3. Termination is shown by (#).

Figure 4 is a representation showing alignment of amino acid sequences encoded by the *nef* genes of HIV-1 NL4-3, D36 PBMC, C18 HIV<sub>StV</sub>, C18 HIV<sub>MBC</sub> and C98 HIV. In  
30 phase termination codons are shown by (#). Identical amino acids are shown by (\*). Residues underlined are those immediately before a deletion.

- 41 -

Figure 5 shows a duplication of NFkB and Sp1 sequences in D36 PBMC, C18 HIV<sub>StV</sub>, C18 HIV<sub>MBC</sub> and C98 HIV demonstrated by alignment of their low homology region sequences with the NFkB-Sp1 region of HIV-1 NL4-3. Nucleotide numbering according to Figure 1. Identity with NL4-3 sequence shown by (\*) and NFkB and Sp1 sites in NL4-3 overlined. Position of *nef*/LTR region sequence deletion shown by (Δ).

Figure 6 is a graphical representation showing replication of C18 and C98 viral isolates and D36 PBMCs from asymptomatic patients in PHA-stimulated PBMCs.

10

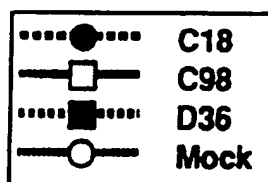
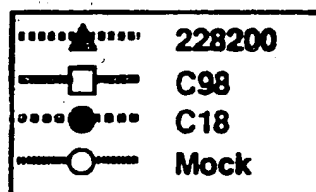


Figure 7 is a graphical representation showing replication of viral isolates from asymptomatic patients in non-PHA stimulated PBMCs.

15

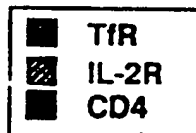


20

228200 is an Australian isolate of HIV-1 described by Kiernan, R. *et al* (1990). Its characteristics include being T cell trophic, with fast kinetics, high producer of HIV-1 and/or SI phenotype.

Figure 8 is a graphical representation of cell surface receptor expression for syncytia-inducing (SI)/ non-syncytia-inducing (NSI)/asymptomatic patient isolates.

30



228200 is defined in the legend to Figure 7. 243925 is a viral isolate of HIV-1 which is monocyte/macrophage trophic and exhibits NSI phenotype (Dr Karen Coats-Fryer, PhD thesis entitled "Viral determinants of HIV-1 syncytium formation", the University of Melbourne, Parkville, Victoria, Australia.

Figure 9 is a representation of the nucleotide sequence of C18 HIV-1<sub>MBC</sub> (SEQ ID NO: 800).

- 10 Figure 10(a) - (g) are graphical representations showing clinical immunology of cohort; (a) CD3; (b)(i) CD4 (ii) CD4%; (c)(i) CD8; (ii) CD8%; (d) lymphocyte count; (e) CD4/CD8 ratio; (f)  $\beta$ -2-microglobulin; and (g) Kaplan-Meier estimates of time to disease progression (Cox & Oakes, 1989).
- 15 Figure 11 is a schematic representation of the deletion mutants of the present invention.



A summary of the SEQ ID Nos. used in the subject specification is shown below:

5	SEQ ID NO:	DESCRIPTION
	1	Nucleotide sequence of HIV-1 strain NL4-3 genome
	2-613	Decanucleotides of <i>nef</i> gene from HIV-1 strain NL4-3
10	614	Partial nucleotide sequence of D36 HIV-1 isolate
	615	Partial nucleotide sequence of C18 HIV-1 <sub>MBC</sub> isolate
	616-625	PCR primers shown in Table 1
	626-633	Sequence primers shown in Table 2
	634	Amino acid residues 15-27 of HIV-1 NL4-3 <i>nef</i> protein
15	635	HIV-1 NL4-3 <i>tat</i> exons (Figure 2)
	636	HIV-1 D36 <i>tat</i> exons (Figure 2)
	637	HIV-1 C18 <i>tat</i> exons (Figure 2)
	638	HIV-1 NL4-3 <i>rev</i> exons (Figure 2)
	639	HIV-1 D36 <i>rev</i> exons (Figure 2)
20	640	HIV-1 C18 <i>rev</i> exons (Figure 2)
	641	HIV-1 NL4-3 C-terminal of gp41 (Figure 3)
	642	HIV-1 D36 C-terminal of gp41 (Figure 3)
	643	HIV-1 C18 C-terminal of gp41 (Figure 3)
	644	HIV-1 NL4-3 <i>nef</i> gene (Figure 4)
25	645	HIV-1 D36 <i>nef</i> gene (Figure 4)
	646	HIV-1 C18 <i>nef</i> gene (Figure 4)
	647	HIV-1 NL4-3 NFkB/SP1 sequence (Figure 5)
	648	HIV-1 D36 NFkB/SP1 sequence (Figure 5)
	649	HIV-1 C18 NFkB/SP1 sequence (Figure 5)
30	650	Nucleotide sequence of <i>nef</i> gene from HIV-1 strain NL4-3
	651	Nucleotide sequence of <i>env</i> and <i>nef</i> regions of NL4-3
	652-799	Decanucleotides of LTR region from HIV-1 strain NL4-3
	800	Nucleotide sequence of C18 HIV-1 <sub>MBC</sub>

**EXAMPLE 1***Source Material*

For the purposes of the following examples, a non-pathogenic HIV-1 strain was isolated  
5 from a recipient of HIV-1 infected blood. The recipient is designated "C18". Other  
recipients are defined as "C54" and "C98". The donor is identified herein as "D36".  
The place of isolation may be indicated after the abbreviation of "HIV". For example,  
St Vincents Hospital, Sydney (HIV<sub>SHV</sub>) or Macfarlane Burnet Centre of Medical  
Research, Melbourne (HIV<sub>MBC</sub>).

10

Exemplary viral isolates referred to herein as "C18" and "C98" were deposited at the  
PHLS Centre for Applied Microbiology and Research, European Collection of Animal  
Cell Cultures (ECACC), Division of Biologies, Porton Down, Salisbury, Wiltshire SP4  
OJG. C18 was deposited on 17 October, 1994 under Provisional Accession Number  
15 V94101706 and C98 was deposited on 31 October, 1994 under Provisional Accession  
Number V941031169.

Figure 11 is a summary of the deletion mutants of the present invention.

20 Viruses were isolated by the following procedures:

1. Infected peripheral blood mononuclear cells (PBMCs) were co-cultured with  
HIV-1 seronegative donor PBMCs. A convenient source of seronegative donor  
PBMCs is a blood bank. The supernatants and cells are harvested every 7 days  
and fresh medium added with CD8 depleted PBMCs. CD8 depletion promotes  
25 the ability to isolate HIV-1. The culture and procedure is continued for up to  
approximately 5 weeks;
2. The infected PBMCs are purified from whole blood and these cells are cultured  
alone for up to 5 weeks. PBMCs alone are used because the virus is more likely  
30 to be monocytotropic. Fresh medium is added weekly and supernatant fluid is  
harvested at this time;

- 45 -

3. Supernatant fluids are harvested every approximately 7 days, fresh medium and fresh HIV-1 seronegative CD8 depleted PBMCs are added at this time;
4. HIV-1 seronegative PBMCs are pretreated with M-CSF for approximately  
5 72 hours prior to the addition of infected PBMCs. M-CSF has been shown to enhance HIV-1 replication in monocytes (Gendelman *et al*, 1988); or
5. The supernatant fluid is harvested from the cultures of step 4 every  
10 approximately 7 days, fresh medium added together with HIV-1 seronegative stimulated CD8 depleted PBMCs. The virus is isolated from the infected PBMCs.

## EXAMPLE 2

### *DNA Preparation and PCR Amplification*

- 15 Non-pathogenic HIV-1 (e.g. strain C18) infected peripheral blood mononuclear cells (PBMC) were harvested 4 days after infection of phytohaemagglutinin (PHA) stimulated HIV-1 negative donor PBMC cultured by the method of Neate *et al* (1987) and washed in phosphate buffered saline (PBS). PBMC from Donor D36 and Recipients C18, C54 and C98 were prepared by Ficoll isopaque centrifugation of buffy coat cells and washed  
20 with PBS.

Approximately  $10^7$  cells were lysed in 1ml lysis solution (0.45% v/v NP40, 0.45% v/v Tween 20, 10mM Tris-HCl pH 8.3, 40mM KCl 2.5 mM  $MgCl_2$ ) and digested with 60µg/ml proteinase K (Boehringer Mannheim) at 55°C for 1 hour followed by 100°C  
25 for 10 minutes. Lysates were stored at -20°C.

All polymerase chain reaction (PCR) primers (Table 1) and sequencing primers (Table 2) were synthesised using an Applied Biosystems model 391 DNA synthesiser using phosphoramidite chemistry.

30

Strict physical separation was maintained for sample, PCR reagent mix and PCR reaction preparations as well as amplification and analysis. Final reaction mixes (50µl) contained

- 46 -

2 µl neat or diluted cell lysate, 0.2µM each primer, 200mM dNTPs and 1.25 units Taq polymerase (Boehringer Mannheim) in PCR buffer, (10mM Tris-HCl pH8.3, 50mM KCl, 100µg/ml gelatine) adjusted to the optimum MgCl<sub>2</sub> concentration for the primer pair (1.5-3.0mM). Aliquoted reagent mix was overlaid with 50µl mineral oil prior to  
5 addition of DNA template lysate. After template denaturation at 94°C for 3 min amplification was achieved with 30 cycles of 94°C, 1 min; 55°C, 1 min; 72°C, 2 mins. A final elongation reaction was conducted at 72°C for 7 minutes. For double PCR amplification 2 µl of first round product was added to the second reagent mix and amplified as before.

10

PCR amplified DNA was checked for quality, quantity and fragment size by agarose gel electrophoresis in Tris-Acetate-EDTA buffer (Sambrook et al, 1989) stained in ethidium bromide and viewed by UV transillumination.

15

### EXAMPLE 3

#### *DNA Sequence Analysis*

The DNA sequence of PCR amplified HIV-1 regions was determined by the dideoxynucleotide method (Sanger et al, 1997) using Sequenase T7 polymerase (United States Biochemicals).

20

PCR amplified DNA was purified by PCR Magic prep resin chromatography (Promega). Approximately 2 to 7 µg purified DNA plus 10ng specific primer (Table 2) were denatured by boiling for 3 mins and snap frozen to -20°C. The initial labelling reaction was for 3 minutes at 22°C (room temperature) with <sup>35</sup>SdNTP (500Ci/mmol; Dupont)  
25 followed by dideoxynucleotide termination reactions at 37°C for 5 minutes. NP40, to 0.45% v/v, was included in denaturation and reaction mixes (Bachman et al, 1990). Sequencing reaction products were denatured in formamide and resolved on a 6% w/v polyacrylamide gel containing 8M urea, fixed in 10% v/v acetic acid, 10% v/v methanol and dried. Following autoradiography on XK1 film (Kodak) the gel sequences were  
30 read assembled, translated to protein and aligned using the PC/GENE suite of programs (IntelliGenetics, USA).

- 47 -

TABLE 1  
PCR PRIMERS

	PRIMER	SEQUENCE <sup>1, 3</sup>	POSITION <sup>2</sup>
5	CI-1	TGGAAGGGCTAATTTGGT(616)	1-18
	CI-2	ATCTTCCCTAAAAAATTAGCCTGTC(617)	2099-2075
	LTR-3'	AGGCTCAGATCTGGTCTAAC(618)	9559-9540
	SK68	AGCAGCAGGAAGCACTATGG(619)	7786-7805
	CI-6	TGCTAGAGATTTTCCACAC(620)	9709-9691
10	KS-2	AGTGAATAGAGTTAGGCAGG(621)	8326-8345
	RT5'-v3	GTAAGACAGTATGATCAGATA(622)	2418-2438
	RT3'-v2	TTGTAGGGAATTCCAAATTCC(623)	4660-4640
	RT5'-v2	CAGGATCCTACACCTGTCAACATAAT(624)	2487-2506
15	RT3'-v1	GGGAATTCCTTATTCCTGCTTG(625)	4655-4634

1. Sequence is presented from 5' to 3' of the primer.
2. Position is according to the numbering of HIV NL4-3 in Myers *et al* (1992).
3. SEQ ID NOs are given in parentheses.

TABLE 2  
SEQUENCING PRIMERS

PRIMER	SEQUENCE <sup>1, 3</sup>	POSITION <sup>2</sup>
KS3	CCAGAAGTTCCACAATCC(626)	8570-8553
KS4	TTCTTCTAGGTATGTGGAG(627)	8753-8735
KS5	AGTGAATTAGCCCTTCCAG(628)	9093-9075
KS6	TGCTAGAGATTTTCCACAC(629)	9709-9691
SP2	TGCTCTGGAAACTCAT(630)	8006-8022
SP3	CTTCTATAGTGAATAGAG(631)	8318-8336
SP4	TATTGGAGTCAGGAAGT(632)	8618-8634
SPR	GGTCTAACCAGAGAGAC(633)	9547-9531

- Sequence is presented from 5' to 3' of the primer.
- Position is according to the numbering of HIV NL4-3 in Myers *et al* (1992).
- SEQ ID NOs are given in parentheses.

#### EXAMPLE 4

##### *Cells and Cell Culture*

Peripheral blood was obtained from HIV-1 sero-negative volunteers and mononuclear cells prepared by centrifugation on a Ficoll/Hypaque density gradient (Peper *et al*, 1968). PBMC were activated with phytohemagglutinin (PHA; 10 $\mu$ g/10<sup>6</sup> cells) for 48 h at 37°C washed and then cultured in RPMI 1640 medium containing 10% v/v heat inactivated foetal calf serum, 15mM HEPES, 0.1% v/v sodium bicarbonate, 25 $\mu$ g/ml polybrene (Sigma), 10% v/v interleukin 2 (Boehringer Mannheim) and 1:1000 anti-interferon (Miles) (IL-2 medium). Non-PHA stimulated cells were prepared in a similar manner except they were cultured in medium lacking PHA and IL-2.

**EXAMPLE 5***Antipeptide-antisera*

Antibodies specific for HIV-1 Nef were raised against a peptide corresponding to the predicted amino acid residues 15-27 (AVRERMRAEPAA SEQ ID NO: 634) of Nef encoded by the HIV-1 clone pNL4.3 (Kemp et al, 1988). The peptide was conjugated to keyhole limpet hemocyanin (KLH; Calbiochem, Behring Diagnostics, CA) via glutaraldehyde and this complex used to immunise sheep (0.5mg peptide conjugate/sheep). Antibodies to the peptide were purified by affinity chromatography. Reactivity of the antibodies with recombinant HIV-1 Nef 25 and 27 was demonstrated by immunoblotting.

**EXAMPLE 6***Reactivity of anti-Nef<sub>(15-27)</sub> with HIV C18-infected Cells in Immunoblotting*

Seven days post-infection HIV-1 C18-infected PBMCs and mock - infected cells were washed in PBS then lysed (0.5% w/v NP-40, 0.5% w/v sodium deoxycholate, 50mM NaCl, 25mM Tris-HCl, 10mM EDTA, 0.01% w/v sodium azide and 10mM phenylmethylsulphonylfluoride). After nuclei were spun out lysates were electrophoresed in a 13% w/v SDS-polyacrylamide gel (SDS-PAGE) and subsequently transferred to Hybond-C nitrocellulose (Amersham, Buckinghamshire, England) for 1 h at 100 V using a Bio-Rad protein transfer cell (Bio-Rad, Richmond, Ca). Membranes were pre-incubated with 1% w/v BSA/PBS for 2 h at room temperature and then reacted with affinity purified sheep anti-Nef<sub>(15-27)</sub>, diluted 1:100, overnight at room temperature. After three washes in 1% w/v BSA/PBS, the blots were incubated with donkey anti-sheep Ig conjugated to biotin (Amersham, diluted 1:500) for 1 h at room temperature. After extensive washing as described above the membranes were incubated with streptavidin-conjugated horse radish peroxidase (Amersham; diluted 1:500 for 1 h at room temperature. All dilutions were made with 1% w/v BSA in PBS. After further washing the membrane was developed with phenylenediamine substrate (Dako, Dapopatts, Denmark). The antibody preparation used in the immunoblotting experiments was free of detectable antibodies to the immunogenic carrier protein and coupling reagent.

- 50 -

**EXAMPLE 7***Analysis by Polymerase Chain Reaction Amplification*

A 5' fragment defined by primers Cl-1 and Cl-2 containing the 5' LTR and part of the *gag* gene was amplified. DNA from HIV-1 C18 infected PBMC gave an amplified  
5 fragment (amplimer) of about 1.9 kb compared with 2.1kb for pHXB2 control template, implying a deletion of about 200bp from HIV C18. Further amplification of this fragment with primers defining the U3 region of the LTR (Cl-1 and LTR-3') gave amplimers of about 300 bp for HIV-1 C18 infected PBMC DNA compared with 340bp for C18 and D36 PBMC DNA and 484 bp for pHXB2 control. This implies the loss of  
10 approximately 140 to 180bp from the U3 region of these proviral DNAs.

To analyse the *nef*-gene-3' LTR region, the nested primer pairs SK68-Cl-6 and KS-2-LTR-3' were used in a double PCR. Amplimers of approximately 830bp were obtained for HIV-1 C18 infected PBMC DNA as well as for PBMC DNA from Donor D36 and  
15 Recipients C18, C54 and C98 compared with approximately 1230bp for pHXB2 DNA. These results suggest that about 400bp of DNA have been deleted from the Donor and Recipient proviral DNAs.

In comparison, amplification of the polymerase gene region by double PCR with the  
20 nested primer pairs RT5'-v3-RT3'-v2 and RT5'-v2-RT3'-v1 gave a fragment (approximately 2.1 kb) the same size as the molecular clone pHXB2 fragment for HIV-1 C18 infected PBMC DNA, suggesting that deletions from this region were unlikely.

**EXAMPLE 8***Nucleotide Sequence of the *nef*-3' LTR Region*

25 PCR amplification experiments indicated an approximately 200bp nucleotide deletion from both the *nef* gene and LTR regions of Donor D36 PBMC and Recipient C18 HIV-1 proviral DNA. To further analyse these regions, the DNA sequence was determined for the PCR amplified *nef*-3' LTR region of D36 PBMC, C18 isolates HIV<sub>MBC</sub> and HIV<sub>Stv</sub>  
30 as well as isolate C98 HIV infected PBMC proviral DNA. The 3' region was amplified with outer primers (SK68-Cl6) and inner primers (SK68-LTR 3' or KS2-Cl6) and sequenced directly using a number of internal sequencing primers based on the HIV-1



NL4-3 nucleotide sequence (Table 2).

Alignment of the nucleotide sequences of the amplified 3' region of donor D36 PBMC and recipient C18 isolates HIV<sub>MBC</sub> and HIV<sub>StV</sub> and C98 HIV (Fig 1) showed a number  
 5 of nucleotide sequences changes, including deletions, relative to the nucleotide sequence of wild-type infectious HIV-1 (HIV-1 NL4-3). In the region of alignment, D36 PBMC lacked 291 nucleotides, C18 HIV<sub>StV</sub> differed in size by 388 nucleotides (comprising deletions of 397 nucleotides and an insertion of 9 nucleotides), C18 HIV<sub>MBC</sub> differed by 456 nucleotides and C98 HIV lacked 158 nucleotides compared with HIV-1 NL4-3.  
 10 The overall identity with HIV-1 NL4-3 nucleotide sequence of D36 PBMC, C18 HIV<sub>StV</sub>, HIV<sub>MBC</sub> and C98 HIV nucleotide sequences, including deletions, was 73% (1157/1596), 67% (1459/1592), 62% (982/1592) and 79% (1105/1399), respectively.

The D36 PBMC sequence differed from HIV-1 NL4-3 in a number of features. A  
 15 change in the wild type *tat* termination codon from TAG to TCG (Ser) extended the third *tat* exon (which starts at splice acceptor 10) by a further 15 amino acids to terminate at a conserved TAG (Fig 1). The resulting C-terminal peptide is rich in charged amino acids (8/15) (Fig 2a). The wild type *rev* termination codon has also changed (TAG to GAG, Glu) and the third *rev* exon is extended for 14 codons to  
 20 terminate at a conserved TAG (Fig 2b). The encoded extra amino acids are mainly polar (11/14) and charged in nature (Fig 2b). The sequence also encodes the C-terminal 237 amino acids of Env gp41 (Fig 3) terminating at the normal termination codon. The D36 PBMC Env amino acid sequence has 85% identity with the HIV NL4-3 sequence, increasing to 89% if similarities are included.

25

There are significant differences from HIV-1 NL4-3 downstream of the *env* (gp41) gene. A change in the fifth *nef* codon, from TGG (Trp) to TGA (Fig 1), introduces an early termination in the D36 PBMC *nef* gene. The encoded Nef protein is identical to the N-terminal 4 amino acids of NL4-3 Nef (Fig 4). Following the early termination there are  
 30 deletions of 33, 47, 93 and 91 nucleotides and a region of low sequence homology, compared with HIV NL4-3, prior to the wild type *nef* termination codon site (HIV NL4-3 nts 9405-9407). As well as removing a significant part of the *nef* gene, these deletions

- 52 -

also bring into phase a further 6 termination codons. While the polypurine tract (plus strand primer binding site) and the first 38 nucleotides of the LTR U3 region are perfectly conserved, downstream U3 region sequences are disrupted by the 93 and 91 nucleotide deletions and the low homology region. The resulting U3 region lacks  
5 recognition sequences for the transcription factors *c-myb*, USF and TCF1 $\alpha$  as well as one of the suggested NF-AT sites (Gaynor et al, 1992). Downstream from the 91 nucleotide deletion, a 59 nucleotide region of low homology contains two extra NFkB enhancer sites 19 nucleotides upstream of the usual site of a pair of NFkB sites, the upstream one of which is altered in its 5'-half in D36. Sequences further downstream  
10 are highly conserved with respect to HIV-1 NL4-3, including the position and number of Sp1 basal promoter sites, TATA box, TAR and polyadenylation signal sequences.

Similar to D36 PBMC, the C18 HIV<sub>STV</sub> and HIV<sub>MBC</sub> sequences show the *tat* third exon to be extended by 15 codons. All but two codons (altered by point mutations) are  
15 identical to those of D36 PBMC (Fig 2a). The *rev* third exon of both C18 isolates is also extended (Fig 2b) but by only three codons, identical to the first three codons of the D36 PBMC *rev* extension. The same 237 amino acid coding region of Env gp41 is found in both the C18 HIV DNA sequences (Fig 3) and shows 85% identity, increasing to 88% if similarities are included, with the same region of the NL4-3 Env gp41.

20 It is in the *nef* gene and LTR regions that the major differences from wild-type HIV-1 arise, just as in D36 PBMC. The *nef* gene of C18 HIV<sub>STV</sub> encodes 24 amino acids with 9 of the 10 N-terminal being identical to the NL4-3 Nef protein (Fig 4). Thereafter, deletions of 177 and 11 nucleotides cause a frameshift and termination at the 25th codon  
25 (Fig 1). Downstream deletions of 120, 82 and 7 nucleotides cause further loss of wild type *nef* gene sequence and bring into phase a further three termination codons.

The *nef* gene of C18 HIV<sub>MBC</sub> encodes only 7 amino acids with only the initiator methionine identical to the NL4-3 Nef protein. This loss of identity and early  
30 termination is brought about by a 250 nucleotide deletion after the fifth nucleotide of the *nef* gene. Downstream deletions of 120 and 86 nucleotides cause further loss of wild-type *nef* gene sequences. In both C18 isolates there is perfect conservation of the

polypurine tract and 29/31 nucleotides at the 5' end of the U3 region immediately before the 120 nucleotide deletion (Fig 1). This deletion together with the downstream 82 and 7 nucleotide deletions in HIV<sub>stV</sub> and 86 nucleotide deletion in HIV<sub>MBC</sub> and the low homology region cause the loss of the 5' half of the NRT-1 site (Yamamoto et al 1992) and the downstream NFAT site. A third NFkB site is present 31 (HIV<sub>stV</sub>) and 33 (HIV<sub>MBC</sub>) nucleotides upstream of the expected pair of NFkB sites which are themselves separated by 13 nucleotides instead of the 4 nucleotides in HIV-1 NL4-3. The 5'-most Sp1 site sequence is slightly altered but sequences downstream including the other 2 Sp1 sites, the TATA box, TAR and polyadenylation signal sequences are identical to HIV-1 NL4-3 sequence.

The three sequences, D36 PBMC, C18 HIV<sub>stV</sub> and C18 HIV<sub>MBC</sub> show a number of similarities consistent with the transmission of virus from person D36 to person C18 as well as a number of differences indicating post-transmission divergence of sequence. All three have *tat* open reading frames (ORFs) extended by 15 codons. All three have extended *rev* ORFs. The new *rev* termination codon in both C18 HIV-1 isolates, three codons downstream of the NL4-3 *rev* termination codon, has a point mutation in D36 PBMC to make a Glu codon so that it continues for a further 11 codons (Fig 2b) to terminate at a conserved TGA. The partial Env gp41 amino acid sequences are more closely related to each other (86% identity or 90% including similarities) than to HIV NL4-3 (85% and 89%, respectively).

The nucleotide sequence of the *nef* and LTR region of the HIV-1 isolate from recipient C98 (C98 HIV) is 90.3% identical (1264/1399) to the HIV NL4-3 sequence, ignoring deletions. Similar to the D36 PBMC and C18 HIV<sub>stV</sub> and HIV<sub>MBC</sub> isolates the C98 HIV sequence shows the third exon of *tat* to be extended by 15 codons with all but one being identical to the D36 PBMC *tat* extension. Also, the *rev* gene is extended by 3 codons, 2 of which are identical to the first 2 codons of the D36 PBMC *rev* extension. The sequence also encodes the C-terminal 223 amino acids of Env gp41 terminating at the normal termination codon. The C98 HIV Env amino acid sequence has 89% identity with HIV NL4-3 Env sequence, increasing to 92% of similarities are included.

- As with the D36 PBMC and the C98 HIV isolate sequences it is the *nef* gene and LTR regions that major differences from the HIV NL4-3 sequence arise. The *nef* gene open reading frame of C98 HIV is much longer than in D36 PBMC, C18 HIV<sub>StV</sub> and HIV<sub>MBC</sub>, encoding 85 amino acids compared with 206 amino acids for NL4-3. Sixty
- 5 eight of those 85 amino acids are identical to the N-terminal sequence of NL4-3 Nef. The single, small deletion (16 nucleotides) in the C98 HIV *nef*-alone regions (Table 3) occurs after *nef* codon 82 causing a frameshift and termination after a further 3 codons at the start of the highly conserved polypurine tract sequence immediately before the 3'-LTR. The *nef*/LTR region has two deletions totalling 142 nucleotides. The 5'-most
- 10 deletion of 42 nucleotides includes the splice acceptor 12 sequence. The NRT-1, dyad symmetry and *myb* response element sequences are all intact. However, the downstream 100 nucleotide deletion includes sequences from the 3' end of the 5'-NF-AT and all of the 3' NF-AT sequences as well as the USF transcription factor recognition sequence. The downstream low homology region of 77 nucleotides lacks the TCF-1 $\alpha$  sequence but
- 15 has two additional NF $\kappa$ B sites 13 nucleotides apart and 26 nucleotides upstream of the 3'-half-remnant of the normal 5'-NF $\kappa$ B site. Sequence downstream, including the 3'-NF $\kappa$ B site, the 3 Sp1 sites, TATA box TAR and polyadenylation signal sequences are all highly conserved.
- 20 The main feature of the sequences is the series of deletions, with respect to HIV NL4-3, in the *nef* gene-3'-LTR region. These can be grouped into two regions namely the *nef*-alone region, that part of the *nef* gene upstream of the LTR, and the *nef*/LTR region, where the *nef* gene and LTR U3 regions overlap. The deletions in these regions of each of the sequences start and end at the same or similar positions (Table 3). The deletions
- 25 are larger in C18 HIV<sub>StV</sub> and C18 HIV<sub>MBC</sub> sequences where totals of 397 and 456 nucleotides have been deleted (relative to NL4-3) compared to 291 nucleotides, from D36 and 158 nucleotides from C98 HIV. In the *nef*-alone region the two deletions in C18 HIV<sub>StV</sub> and the single deletion in C18 HIV<sub>MBC</sub> occupy the same region as the three deletions in D36 PBMC. Similarly, the *nef*/LTR region in the three deletions in the C18
- 30 HIV<sub>StV</sub>, the two deletions in the C18 HIV<sub>MBC</sub> and the D36 PBMC sequences occupy the same region. These findings indicate that mutant virus was transmitted from D36 to C18 after which further deletions and rearrangements occurred. Similarly, the

- 55 -

sequence of C98 HIV in the *nef*/LTR region indicates two deletions occupying the same region as the *nef*/LTR deletions in D36 and the C18 sequences. However, the size (only 16 nucleotides) and the position of the deletion in the *nef*-alone region of C98 HIV are distinct from those of the D36 PBMC and C18 sequences.

5

The timing of transmission of virus by transfusion was that recipient C18 was transfused approximately 19 months after C98. Consistent with the relative timing of transmission and the sequence similarities and differences is the suggestion that at the time of transmission to C98, the D36 sequence had deletions in the *nef*/LTR region but not in  
10 *nef*-alone region. After transmission to C98, the C98 virus developed further deletions and rearrangements, including the deletion in the *nef*-alone region. The D36 virus evolved so that at the time of transmission to C18, further deletions and rearrangements had occurred including deletion of sequence from the *nef*-alone region distinct from the C98 HIV *nef*-alone region deletion. After transmission to C18, further deletions and  
15 rearrangements occurred in the C18 virus giving rise to at least two sequences (HIV<sub>Stv</sub> and HIV<sub>MBC</sub>).

The *nef*-alone deletion region may be a mutation or recombination "hotspot" as it includes sequences that were found to be variably duplicated in 28 out of 54 Nef protein  
20 sequences derived from 8 of 12 patients analysed in a study (Shugars *et al* 1993). The sequence between the *nef*-alone and the *nef*/LTR region deletions is highly conserved and is important in provirus integration into the infected cell genome and interacts with a number of cellular proteins. It is interesting that the sequence equivalent to NL4-3 nucleotides 9209 to 9225 is retained in D36 and C98 HIV but lost in the C18 HIV  
25 sequences. This includes part of a sequence of dyad symmetry (9210 to 9231) and is a significant part of the binding site for NRT-1 (Yamamoto *et al* 1991) which has been shown to have a negative regulatory effect on HIV-1 expression. The presence of this sequence in D36 and C98 HIV and its absence from the C18 isolates may correlate with the inability to isolate virus from D36 PBMC and the poor replication of C98 HIV but  
30 the ability to isolate HIV-1 from C18 PBMC. The deletion of sequence equivalent to nucleotides 9281 to 9395 of NL4-3 causes the loss of some transcription factor binding sites including NFAT and USF from the D36, C18 HIV and C98 HIV sequences.

- 56 -

A further similarity between the D36, C18 HIV<sub>StV</sub>, C18 HIV<sub>MBC</sub> and C98 HIV sequences is a region of low homology to HIV-1 NL4-3 extending downstream of the *nef*/LTR deleted region to the NFkB enhancer/Sp1 promoter site region. This low homology region in fact consists of incomplete duplications of part of the NFkB/Sp1  
5 region (Fig 5) resulting in D36 and C98 HIV having 2 extra NFkB sites upstream of an altered 5' NFkB site while the C18 sequences have one extra NFkB site and altered spacing between the 5' and 3' wild type NFkB sites due to an insertion of 9 nucleotides.

For the C18 and C98 HIV-1 isolates virus replication was assessed in PHA-stimulated  
10 and non-stimulated PBMCs (Figs 6 and 7). In PHA-stimulated PBMCs we also studied cell surface CD4 and IL-2R expression (Fig 8). In comparison with HIV-1 wild type SI and NSI isolates clearly both C18 HIV<sub>MBC</sub> and C98 viruses are replication competent, though C98 HIV replicates more poorly than C18 HIV<sub>MBC</sub> and are of the NSI phenotype when syncytium formation and CD4 and IL-2R surface expression are  
15 taken into account. Additionally, and more surprisingly, these two viruses replicated almost as efficiently in non-PHA stimulated PBMCs when compared to a typical local wild type SI isolate (HIV-1 228200, Fig 7).

When protein expression was assessed for C18 HIV<sub>MBC</sub> and C98 HIV<sub>MBC</sub> although  
20 structural proteins were identified, no typical Nef protein was seen in infected cells. However, analysis of cell lysates prepared from PBMC infected with C18 HIV<sub>MBC</sub> or PBMC infected with C98 HIV<sub>MBC</sub> (which were subsequently stimulated by UV irradiation, see Valerie *et al*, 1988) by Western immunoblotting using two antibodies specific for the N-terminal region of Nef showed the presence of smaller proteins of 19  
25 kDa and 21 kDa, respectively. These proteins were not observed in mock-injected control PBMC lysates and were not observed when the infected-cell lysates were probed were probed with antibodies reactive with the C-terminus of Nef.

Thus, although the C18 and C98 HIV isolates are replication competent *in vitro* they  
30 clearly replicate differently using different conditions for cell activation and from the known functions of HIV-1 Nef protein and the LTR show that the major deletion in the *nef* gene and/or the LTR is at least in part responsible for the outcome of infection,

- 57 -

implicating the importance of Nef and/or the LTR in the clinical outcome of infection *in vivo*.

#### EXAMPLE 9

##### 5                    *Determination of Degree of Relatedness Between Viruses*

To determine the degree of relatedness between viruses such as between mutants or between mutants and a wild-type virus and to ascertain putative infected patients, the method of Delwart *et al* was employed.

#### 10                    EXAMPLE 10

##### *Immune responsiveness of subjects infected by non-pathogenic HIV-1 isolate*

In this example, the donor and recipients of the cohort were tissue typed and assessed for basic cellular immune responses. Proliferative responses and IL-2 production to the mitogens ConA and PHA, to allogeneic mononuclear cells (irradiated pooled  
15 mononuclear cells from 20 random donors) and to recall antigens (e.g. influenza and tetanus toxoid) were within normal ranges. While at the immunogenetic level, HLA typing failed to identify a consistently common allele or haplotype within the group.

The conservation of CD4+ counts observed in the cohort, the relative integrity of their  
20 immune system and the varied HLA types of the donor and recipients further supports the fact that the symptomless condition of the cohort members is due to a non-pathogenic strain of HIV-1 or a strain of low virulence.

Accordingly, this provides a screening procedure for subjects putatively infected by a  
25 non-pathogenic HIV-1 isolate where such subjects are seropositive for HIV-1 (e.g. have antibodies to an HIV-1 glycoprotein) yet have normal proliferative responses and cytokine production to mitogens, allogeneic mononuclear cells and to recall antigens.

**EXAMPLE 11***Clinical Immunology of Cohort*

To establish that the donor and the recipients belonging to the cohort exhibit normal immunological profiles, members of the cohort were assayed for CD3, CD4, CD8, lymphocyte count, CD4/CD8 ratio and  $\beta$ -2-microglobulin over time since seroconversion.

Parameters considered normal in non-infected individuals range as follows:

10

<u>Parameter</u>		
CD3	55-82%	620-2200 ( $\times 10^6/L$ )
CD4	29-58%	420-1410 ( $\times 10^6/L$ )
CD8	12-43%	200-980 ( $\times 10^6/L$ )
15 Lymphocyte count	1000-3500 ( $\times 10^6/L$ )	
CD4/CD8	0.7-3.7	
$\beta$ -2-microglobulin	0.00-2.20 mg/L	

20 The results are shown in Figures 10(a)-(g) and clearly show that the immunological profiles of cohort members are substantially normal further highlighting the non-pathogenicity of the HIV-1 isolates of the present invention. Figure 10(g) shows a graph of the Kaplan-Meier (Ox and Oates, 1989) estimates of time to disease progression (AIDS or  $CD4 > 250$ ). The results demonstrate that the difference is large in spite of

25 the conservative bias, with a median time to progression of 6.2 years in the main database. An exact logrank test (Cytel Software Corporation, 1989, StatXact: Statistical Software for Exact Nonparametric Inference. Cambridge, Massachusetts.) was performed, demonstrating that the difference between the groups was highly statistically significant (logrank statistic 11.8,  $p < 0.0001$ ).

30



**Table 3**  
**Deletions and their sizes in the *nef*-alone and the *nef*/LTR regions of the Long-Term Asymptomatic HIV-1 Sequences**

Sequence	<i>nef</i> -alone Region	Region Deletion (nt)	<i>nef</i> /LTR Region	Region Deletion (nt)	Total Deletion (nt)
D36 PBMC	8830-8862 (33)		9112-9204 (93)		
	8882-8928 (47)		9281-9371 (91)	184	291
	9009-9035 (27)	107			
C18 HIV <sub>StV</sub>	8830-9006 (177)		9105-9224 (120)		
	9019-9029 (11)	188	9281-9362 (82)	202	390
C18 HIV <sub>MBC</sub>	8792-9041 (250)	250	9105-9224 (120)		
			9281-9366 (86)	206	456
C98 HIV	9033-9048 (16)	16	9148-9189 (42)		
			9271-9370 (100)	142	158
C54 PBMC	incomplete	?	9281-9375 (95)	95	95 +

Sequence numbering relates to the equivalent position in HIV-1 NL4-3. Numbers in brackets are the deletion sizes in nucleotides (nt).  
The *nef* ORF starts at nt 8787 and the 3'-LTR starts at nt 9074 in HIV-1 NL4-3.

## EXAMPLE 12

*Sequencing of isolate HIV-1 C18<sub>MBC</sub>*

The genome of variant HIV-1 designated C18 HIV-1<sub>MBC</sub> was amplified by the  
5 polymerase chain reaction (PCR) as 7 overlapping fragments using the sets of inner and  
outer oligonucleotide primers, designed using the programme PCRPLAN  
(IntelliGenetics), listed in Table 5 and either UITma (Applied Biosystems) or a mixture  
of KlenTaq and Pfu (KlenTaq LA, Ab Peptides Inc) polymerases (for faithful  
amplification of long fragments). The resulting fragments were cloned into the SmaI  
10 site of the plasmid vector pGEM7Zf+. Insert-containing clones representing each region  
of the full length variant HIV-1 were sequenced by a nested deletion strategy (Gou &  
Wu, 1982) and cycle sequencing with Taq polymerase and dye labelled primers  
complimentary to the T7 or SP6 sites within the cloning vector. Nucleotide sequences  
were entered and collated by ASSEMBL and SEQIN (IntelliGenetics) and SEQED  
15 (Applied Biosystems) and translated to the encoded amino acid sequences using  
TRANSL (IntelliGenetics) programmes. Sequence alignments used NALIGN,  
CLUSTAL (IntelliGenetics) and SEQED programmes.

The full length sequence (Fig 9; SEQ ID NO:800) of isolate HIV-1 C18<sub>MBC</sub> is 9207  
20 nucleotides long which is 506 nucleotides shorter than the HIV NL4-3 sequence. This  
size difference is comprised of 126 nucleotides of insertions and 632 nucleotides of  
deletions, see Table 6. The most extensive differences between the HIV-1 C18<sub>MBC</sub>  
sequence and HIV-1 NL4-3 are in the U3 region of the LTR and in the *nef* gene, as  
hereinafter described.

25

The 5' LTR has deletions of 120 and 87 nucleotides and a region of low sequence  
homology, which is the result of an imperfect duplication of the downstream NFκB and  
Sp1 response sequences. These result in the loss of sequence from a number sites  
important in the regulation of transcription of HIV-1 genes, including the negative  
30 response element (NRE) and the response elements for a number of transcription factors  
including NF-AT, NRT-1, USF and TCF-1α. Furthermore, the low homology region  
contains an extra NFκB and Sp1 sites as well as an insertion of 9 nucleotides between

- 61 -

the usual NF $\kappa$ B sites. Downstream of the NF $\kappa$ B sites the sequence of the LTR has a high level of homology (96.2%) with the same region of HIV NL4-3.

The *gag* gene contains 3 insertions, which represent direct repetitions of adjacent sequences. The first is a perfect repeat of 15 nucleotides after the equivalent of nucleotide 1134 of HIV-1 NL4-3 and adds 5 amino acids to the C-terminus region of p17<sup>gag</sup>. The remaining 2 insertions are imperfect and perfect repeats of 30 and 6 nucleotides, respectively, after the equivalent of HIV NL4-3 nucleotides 2163 and 2232, respectively. These encode an extra 12 amino acids in the C-terminus region of p15<sup>gag</sup> just downstream of the *gag* to *pol* frameshift sequences. The variation in sequence length of the *gag* gene at these two positions is unusual. The homology of the encoded amino acid sequence of HIV-1 C18<sub>MBC</sub> and HIV NL4-3 for the *gag* p17, p24, and p15 proteins is 87.1%, 93.5% and 94.3%, respectively.

In the *pol* ORF, the encoded proteins have high homology with the NL4-3 sequences being 95.5% overall comprising p10 protease 92.9%, p66 reverse transcriptase 95.4% and p34 integrase 95.8%. The amino acid sequence of the p61 RT lacks the mutations associated with resistance to the nucleoside (AZT, ddI, ddC) and non-nucleoside (Nevirapine) analogue drugs used in the treatment of HIV-1-infected persons.

20

The *vif* gene encodes a 192 amino acid protein with 88.0% homology with that of HIV-1 NL4-3. The *vpr* gene encodes a 96 amino acid protein with 89.6% homology with that of HIV-1 NL4-3.

There are 2 insertions and 1 deletion of sequences in the *vpu* gene. The insertions of 3 and 9 nucleotides are after the equivalent of nucleotide 6071 and 6234, respectively, of HIV-1 NL4-3. These add 1 amino acid after amino acid 3, and 3 amino acids after amino acid 59 of the encoded *Vpu* protein. The deletion of 12 nucleotides after the equivalent of HIV-1 NL4-3 nucleotide 6261 deletes 4 amino acids from the C-terminal region of *Vpu* as well as from the signal peptide of the *env* polyprotein, which is encoded by an overlapping reading frame. Amino acid sequence homology of HIV-1 C18<sub>MBC</sub> *Vpu* with NL4-3 is 85.2%.

30

- 62 -

The sequence encoding the *env* gp120 has 9 insertions totalling 45 nucleotides (encoding 15 amino acids) and the deletion of a total of 18 nucleotides (encoding 6 amino acids). These are listed in Table 6. All of these events (insertions and deletions) are at positions in the *env* gene. This is within the *env* V3 coding region, immediately upstream of the sequence encoding the so called V3 tip (or loop) amino acid sequence, Gly Pro Gly Arg. The V3 region sequence is that of a typical clade B subtype (North America, Europe and Australia) being identical to the clade B consensus sequence (based on 186 *env* sequences) at 29/35 positions. The type of amino acid at positions 11 and 28 of the V3 loop region (where position 1 is the Cys at amino acid 266 of the *env* gp120) is predictive of the viral non-syncytium / syncytium forming phenotype (Fouchier et al, 1992). The HIV-1 C18<sub>MBC</sub> *env* gene encoded amino acid sequence has Ser at position 11 and Ile at position 28 of the V3 loop region. The lack of a positively charged amino acid at both positions is strongly indicative of a non-syncytium viral phenotype. The overall amino acid sequence homology with HIV-1 NL4-3 (ignoring deletions and insertions) is 86.1%, comprising 85.5% for the gp120 region and 87.6% for the gp41 region.

Both the *tat* and *rev* second exon open reading frames (ORF) are longer than in HIV-1 NL4-3. A change of the *tat* termination codon from TAG to TCG extends the *tat* ORF to a downstream in phase termination codon extending the encoded *tat* amino acid sequence by 15 residues, compared with the 86 amino acid long NL4-3 *tat* protein, to a total length of 101 amino acids. However, this is the usual length of the HIV-1 *tat* protein.

Similarly, the normal *rev* termination codon is changed from TAG to GAG. This extends the *rev* ORF to an in-phase termination codon 3 codons downstream so that the encoded Rev protein is 119 amino acids long instead of the usual 116.

As mentioned above the most extensive differences between the sequences of the isolate HIV-1 C18<sub>MBC</sub> and HIV-1 NL4-3 are in the *nef* gene and the LTR region. While the *nef* gene overlaps the 3' LTR, these differences are found in both the *nef* alone and the *nef* / LTR overlap regions. The HIV C18<sub>MBC</sub>-encoded *nef* protein is only 24 amino

- 63 -

acids long compared with the normal length of 206 amino acids. This severe shortening of the *nef* protein is due to the deletion of 188 nucleotides (the 177 and 11 nucleotide deletions) from the *nef*-alone region which also brings into phase a termination codon, TAG, at the resulting 25th codon. Downstream there is further loss of potential *nef* gene sequences by the 120 and 87 nucleotide deletions situated in the *nef* / LTR overlap region. The resulting 24 amino acid *nef* protein is identical to the N-terminus of the HIV-1 NL4-3 *nef* at 9 of the first 10 positions. Thereafter, homology is lost completely.

Some sequences used in the generation of mature mRNAs are altered or lost in C18<sub>MBC</sub>. The dinucleotide immediately after the splice donor site 2 (SD2) at nts 4818-4819 (NL4-3 equivalent nts 4963-4964) is changed from the conserved GT to GC. It is expected that this change would lead to loss of function of this site as a splice donor. Splice donor 2 is used in the processing of HIV-1 transcripts to some of the mRNAs that encode Tat, Rev and *nef* proteins. Similarly the splice acceptor site 7 (SA7) sequence at nts 6477-6478 (NL4-3 equivalent nts 6602-6603) is changed from the conserved AG dinucleotide to TC. This change is expected to lead to loss of function of this site as a splice acceptor. While this SA site is used in HIV-1 mRNA processing it is not a major site and is not used in the production of the regulatory proteins (Tat, Rev or *nef*) mRNAs. The splice donor 12 site is absent from the C18<sub>MBC</sub> sequence (NL43 equivalent nts 9161-9162) as it is within the first deletion region in the *nef* / LTR overlap region which occurs at nt 8797 and results in the loss of NL43 nucleotides 9105 to 9224. It is significant that the SA12 site is absent from the sequence of all of the cohort virus isolates so far obtained as well as from the sequence of D36 PBMC, however, the C54 PBMC sequence does contain the SA12 site. SA12 is not used in the processing of mRNAs that encode the viral regulatory proteins. Normally SA12 is used in splicing in conjunction with SD1, 2, 3 and 4 and the resulting spliced RNA is probably not a mRNA but may have a regulatory role involving binding to cellular proteins (Smith et al, 1992).

An interesting feature of the sequence of the HIV-1 C18<sub>MBC</sub> isolate is the deletion and rearrangement of sequence from the 5'-LTR U3 region and the deletion of sequence from the *nef* gene (both *nef* alone and *nef* / 3' LTR regions). These being the only

- 64 -

features of the sequence distinct from disease-causing HIV-1. The lack of AIDS or AIDS-like symptoms in the patient C18 is attributed to the effects of the loss of LTR sequence and / or the loss of *nef* coding sequences and their role in the pathogenesis of AIDS.

5

**TABLE 4**  
**Primers used to Amplify Overlapping regions of HIV-1 C18<sub>MBC</sub>**

10	Primer	5'- Coordinate	Direction (+/-)	Primer Length	(nt)	Sequence
15	CL 1A	1	+	30		TGGAAGGGCTAATTTACTCCAAAAAGAC
	CL 14	896	-	25		AATCGTTCTAGCTCCCTGCTTGCCC
	CL 1B	1	+	30		<u>AATCCGGGTGGAAGGGCTAATTTACTCCC</u>
	CL 13	796	-	31		<u>CCTCTAGACCGCTTAATACTGACGCTCTCGC</u>
20	CL 11	682	+	23		TCTCTCGACGCAGGACTCGGCTT
	CL 18	3440	-	30		CTGTTTTCTGCCAGTTCTAGCTCTGCTTCT
	CL 12A	732	+	26		<u>TTTCCCGGGCGGCGACTGGTGAGTAC</u>
25	CL 17	3330	-	32		<u>CCCTCTAGACTTGCCCAATTCAATTTTCCAC</u>
	CL 26	3193	+	39		CCACACCAGACAAAAGCATCAGAAAGAACCCCATTC
	CL 6B	9671	-	39		TGCTAGAGATTTTCCACCGGACTAAATGGTCTGAGGG
30	CL 27	3251	+	39		CCATCCTGATAAATGGACAGTACAACCATAGTACTGCC
	CL 28	639	-	37		TGGCCCAAACATTATGTACCTCTGCATCATATGC
	CL 19	5448	+	30		AGCAGGACATAACAAGGTAGGATCTCTACA
	CL 24	8422	-	28		GGATCTGTCTCTGTCTCTCTCCACCT

35

Underlined sequences depict added restriction enzyme site

+ and - orientations refer to sense and antisense strands of the double stranded DNA sequence

**TABLE 5**  
**Sequence Deletions and Insertions in HIV-1 C18<sub>MBC</sub>**  
**Compared with HIV-1 NL43**

5	Gene or Region	Position (nt)		Deletions (nt)	Insertions (nt)
		C18 <sub>MBC</sub>	NL43		
10	5'-LTR U3	29	29	120	-
	5'-LTR U3	85	205	87	-
	5'-LTR U3	154	360	-	9
	<i>gag</i> p17	939	1134	-	15
15	<i>gag</i> p15	1982	2163	-	30
	<i>gag</i> p15	2081	2232	-	6
	<i>vpu</i>	5927	6062	-	3
	<i>vpu/env</i>	6092	6234	-	9
	<i>vpu/env</i>	6128	6261	12	-
20	<i>env</i>	6483	6628	-	6
	<i>env</i>	6514	6653	2	-
	<i>env</i>	6524	6665	1	-
	<i>env</i>	6630	6772	-	9
	<i>env</i>	6646	6778	-	3
25	<i>env</i>	7011	7141	6	-
	<i>env</i>	7140	7276	3	-
	<i>env</i>	7195	7334	-	6
	<i>env</i>	7266	7399	3	-
	<i>env</i>	7278	7414	-	6
30	<i>env</i>	7290	7420	-	2
	<i>env</i>	7300	7429	-	1
	<i>env</i>	7314	7441	3	-
	<i>env</i>	7463	7593	-	3
	<i>env</i>	7471	7598	-	9
35	<i>nef</i>	8711	8829	177	-
	<i>nef</i>	8723	9018	11	-
	<i>nef</i> / LTR	8798	9104	120	-
	<i>nef</i> / LTR	8854	9280	87	-
	LTR U3	8923	9435	-	9
40				<u>632</u>	<u>126</u>

**EXAMPLE 13***Macrophage Isolates of HIV-1 C18 and HIV-1 C98*

- 5 HIV-1 has been isolated from the macrophages of patients C18 and C98.

Patient monocytes were prepared as follows. Whole blood was spun at 2000rpm for 10 minutes. Plasma was removed into a separate tube and the remaining cells were diluted 1:2 in PBS<sup>-</sup> (magnesium and calcium free phosphate buffered saline). This was  
10 underlaid with 10 ml of Ficoll Isopaque and spun at 2000rpm for 20 minutes. Cells were collected from the interphase and washed three times with PBS<sup>-</sup>. These cells were then seeded into a 6 well Costar tray at a concentration of  $1.0 \times 10^7$ /ml and allowed to adhere for 1 hour. Any non-adherent cells were removed by aspiration.

- 15 Donor HIV-1 negative macrophages for use in co-cultivation were prepared as follows. Peripheral blood mononuclear cells were purified from whole blood using Ficoll/Isopaque density gradient. These cells were seeded at a concentration of  $2.0 \times 10^6$ /ml in teflon. PBMC were cultured in the presence of  $3\mu\text{g/ml}$  of PHA and 1000U/ml of M-CSF 3 days prior to co-culture.

20

- On day of co-culture, donor PBMC were CD8 depleted. Dyna beads coated with anti-CD8 were used for this purpose. Dyna beads were washed once in PBS<sup>-</sup> and then applied to a magnet for 3 minutes. Supernatant was removed and the beads were then resuspended in 250 $\mu\text{l}$  of RF-10. Aliquots of  $2.0 \times 10^8$  patient cells were then added to  
25 250 $\mu\text{l}$  (3 beads : 1 CD8 T-cell) of Dyna beads and allowed to incubate for 30 minutes on ice with occasional mixing. After 30 minutes the cell suspension was placed onto a magnet for 3 minutes. The supernatant was then removed placed into a second tube containing 142  $\mu\text{l}$  (1 bead : 1 CD8 T-cell) of Dyna beads. This suspension was placed on ice for an additional 30 minutes with occasional mixing. After 30 minutes cell  
30 suspension was placed onto a magnet for 3 minutes. Supernatant was removed and washed once in RF-10.



- 67 -

For co-culture, CD8 depleted PBMC were then added to patient monocytes. Half media changes were done every 7 days for a period of 21 days. Aliquots of 2.5 ml of medium was removed from these cultures and replaced with CD8 depleted donor PBMC in Iscoves containing 10% HuS (Human serum), 5% v/v FCS and 5% w/v IL-2 and  
5 1000U/ml of M-CSF. Harvested supernatants were spun at 1400 rpm for 10 minutes and stored as 1ml aliquots. Cell pellets were lysed in 200µl of lysis buffer for PCR analysis. Infection was quantitated using a p24 EIA Kit.

Cells were harvested from the co-cultures and used to prepare DNA as described above.  
10 The *nef* / 3'-LTR region of both virus isolates was amplified by PCR using the above described primer sets and conditions (Example 12). The resulting amplimers were cloned into the plasmid vector pT7T3U19 and the nucleotide sequence determined by the Taq cycle sequencing method with dye-labelled primers.

15 The C18 macrophage sequence has 3 deletions starting and finishing at positions within 3 nucleotides of the same deletions in C18<sub>MBC</sub>. The encoded *nef* protein is 3 amino acids long compared with 7 amino acids for C18<sub>MBC</sub>. The low homology region of the LTR U 3 region of C18 macrophage is very similar in sequence to C18<sub>MBC</sub> and similarly it has one extra upstream NFκB site.

20

On the other hand, the sequence of C98 macrophage has a number of differences from the C98 isolate. While it has exactly the same first deletion of 16 nucleotides just upstream of the polypurine tract (PPT), in the *nef*-alone region, and exactly the same second deletion (position and size) it has an extra deletion of 18 nucleotides at NL4-3  
25 equivalent nucleotides 9206 to 9223. The final deletion is in approximately the same position as in the C98 isolate but is 5 nucleotides longer. The encoded *nef* protein is 34 amino acids long compared with 86 amino acids for the C98 isolate. The low homology region is very similar to the C98 isolate, having the same 2 extra upstream NFκB sites and completely lacking the normal 5'-NFκB site.

30

## EXAMPLE 14

### *Construction and Use of an Infectious Molecular Clone*

5 Molecular biological techniques can be used to construct a molecular clone of, for example, HIV-1 C18<sub>MBC</sub>. Two schemes may be used. In the first scheme genomic DNA, extracted from either the CD4 positive PBMC of the patient C18 or donor PBMC that have been infected with the isolate HIV-1 C18<sub>MBC</sub>, is used as the template for polymerase chain reaction amplification, using thermostable polymerase of high  
10 transcriptional fidelity (eg UITma polymerase or KlenTaq/Pfu polymerase mixture), of long (6 to 7 kb) overlapping fragments representing the 5'- and 3'-parts of the HIV-1 C18<sub>MBC</sub> proviral genome of total length 9207nts. The amplified fragments may then be ligated together after digestion with a restriction enzyme that cleaves at a unique site common to the overlapping region of the amplified fragments, for example the unique  
15 Bgl I or Nco I sites. Ligation of this full length proviral DNA into a plasmid vector will allow its propagation in *E coli* and the subsequent preparation of large (mg) quantities of this molecularly cloned proviral DNA.

In the second scheme donor PBMC that have been infected with the isolate HIV-1  
20 C18<sub>MBC</sub> are used as a source of non-integrated proviral DNA which can be extracted from the infected cells by the Hirt extraction method (Hirt, 1967). Circular proviral DNA molecules may be linearised by digestion with a restriction enzyme that cleaves at a unique position in the genome (eg the Bgl I or Nco I sites). The resulting linearised molecules can be ligated into a plasmid or, more usually, a bacteriophage lambda ( $\lambda$ )  
25 based vector (eg Charon 4a,  $\lambda$ WES) after modification of the end to provide blunt or cohesive ends compatible with the vector. Transformation or transduction of *E coli* with the recombinant plasmid or bacteriophage material, respectively allows the propagation of the proviral DNA. Clones of *E coli* containing proviral DNA may be selected and DNA prepared. Molecular clones of retroviral genomes prepared in this way are often  
30 permuted. Rearrangement to the functional arrangement of sequences is achieved by restriction enzyme cleavage and religation of fragments to reconstruct the correctly permuted proviral genome.

- 69 -

The molecularly cloned DNA products of both schemes can be used to prepare variant proviral genomes that may be used as the basis of a biologically attenuated HIV-1 vaccine strain. Similarly, they may be modified to contain extra DNA sequences in the *nef*-alone deletion region that may deliver sequences that may be of therapeutic  
5 advantage (eg antisense or ribozyme sequences).

Infectious virus particles of HIV-1 C18<sub>MBC</sub>, or modified virus, can be produced by transfection of human cells (eg HeLa cells) which will produce, and release to the culture medium, virus particles of HIV-1 C18<sub>MBC</sub>, or modified virus. These virus  
10 particles can be used to infect a variety of CD4 positive cells for further propagation or experimentation.

### EXAMPLE 15

#### *In vivo Primate Model*

15 Following construction of infectious molecular clones of the mutant HIV-1 strains, studies are then undertaken in primates to establish attenuation, immunogenicity and vaccine prophylactic efficacy. All studies compare mutant clones of HIV-1 with isogenic wild-type (WT) virus. Initial studies are performed using the macaque (*M. nemistrina*) model of HIV-1 infection. Macaque-infectious WT HIV-1 and mutant  
20 clones are compared with respect to duration of viremia, anatomic sites of replication, and cellular and humoral immune responses. Where the mutant HIV-1 clones induce an immune response in the macaques infected, challenge studies with WT virus are also performed. Studies are performed in a limited number of chimpanzees, generally in parallel with the macaque studies. Relevant mutations are engineered into WT HIV-1  
25 clones previously shown to produce chronic infection in chimps, and the course of chimp infection with mutant clones compared with historical controls. If infection is established, WT challenge studies is also performed.

- 70 -

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The  
5 invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

## BIBLIOGRAPHY

- BACHMANN B, LÜKE W AND HUNSMANN G (1990). *Nucl Acids Res* 18: 1309.
- 5
- BARRE-SINOUSSE F, CHERMANN JC, REY F, *et al* (1983) *Science* 220: 868-871.
- BUSCH MP, EL AMAD Z, SHEPPARD HW, ASCHER MS, LANG W (1991) *N Engl J Med* 325: 733-735.
- 10
- CHENG-MAYER C, SHIODA T, LEVY JA (1991) *J Virol* 65: 6931-6941.
- CLERICI M, STOCKS NI, ZAJAC RA, *et al* (1989) *J Clin Invest* 84: 1892-1899.
- 15
- COX and OAKES (1989) *Survival Analysis, Chapman & Hall*.
- DANIEL MD, KIRCHHOFF F, CZAJAK SC, SEHGAL P AND DESROSIERS RC (1992). *Science* 258: 1938-1941.
- 20
- DELWART EL, *et al* (1993) *Science* 262: 1257-1261.
- FOUCHER RAM, *et al* (1992) *J. Virol.* 66: 3183-3187.
- GAYNOR R (1992). *AIDS* 6m 347-363.
- 25
- GOU L-H and WU R. (1982) *Nucleic Acids Research* 10: 2065-2084.
- GREENWAY *et al* (1994) *Virology* 198: 245-256.
- 30
- GROENINK M, FOUCHIER RAM, BROERSEN S, *et al* (1993) *Science* 260: 1513-1515.

- HAMMES SR, DIXON EP, MALIM MH, CULLEN BR and GREENE WC, (1989)  
*Proc Natl Acad Sci USA* 86: 9549-9553.
- HIRT B. (1967) *J Mol Biol* 26: 365-369.
- 5 HWANG SS, BOYLE TJ, LYERLY HK, CULLEN BR (1991) *Science* 253: 71-74.
- KASLOW RA, DUQUESNOY R, VAN RADEN M, KINGSLEY L, MARRARI M.  
(1990) *Lancet* 335: 927-930.
- 10 KEMP BE, RYLATT DB, BUNDESEN PG, DOHERTY RR, MCPHEE DA,  
STAPLETON D, COTTIS LE, WILSON K, JOHN MA, KHAN JM, DINH DP, MILES  
S & HILLYARD CJ (1988). *Science* 241: 1352-1354.
- KESTLER HW, RINGLER DJ, MORI K, PANICALI DL, SEHGAL PK, DANIEL MD  
15 & DESROSIERS RC (1991). *Cell* 65: 651-662.
- KIERNAN R, et al (1990) *AIDS Res. Hum. Retroviruses* 6: 743-752.
- KIM S, IKEUCHI K, BYRN R, GROOPMAN J and BALTIMORE D (1989) *Proc Natl*  
20 *Acad Sci USA* 86: 9544-9548.
- LANG W, PERKINS H, ANDERSON RE, ROYCE R, JEWELL N. WILKELSTEIN  
W. (1989) *J Acquir Immune Defic Syndr* 2: 63-69.
- 25 LEARMONT J, TINDALL B, EVANS L, CUNNINGHAM A, CUNNINGHAM P,  
WELLS J, PENNY R, KALDOR J AND COOPER DA. (1992). *Lancet* 340: 863-867.
- LEGUERN M, SHIODA T, LEVY JA, CHENG-MAYER C. (1993) *Virology* 195: 441-  
447.
- 30 LEVY JA, (1993) *AIDS* 7: 1401-1410.

- 73 -

LIFSON AR, BUCHBINDER SP, SHEPPARD HW, *et al* (1991) *J infect Dis* 163: 959-965.

LUCIW PA, CHENG-MAYER C and LEVY JA (1987) *Proc Natl Acad Sci USA* 84:  
5 1434-1438.

MANIATIS T *et al* (1982). Molecular cloning. A Laboratory Manual, 1st edition Pub. Cold Spring Harbor Laboratory Press.

10 MOSIER DE, GULIZIA RJ, MACISAAC PD, TORBETT BE, LEVY JA (1993) *Science* 260: 689-692.

MYERS A, KORBAR B, BERZOFKY JA, SMITH RF & PAVLAKIS SA eds (1992; 1993). Human retroviruses and AIDS. A compilation and analysis of nucleic acid and  
15 amino acid sequences. Pub Theoretical Biology and Biophysics Group. Los Alamos National Laboratory, Los Alamos, NM USA.

NEATE EV, HEALY DS, PRINGLE RC, GUST ID, AND JOWLETT JMB (1987). *Aust NZ J Med* 17: 461-466.

20

NIEDERMAN TMJ, THIELAN BJ, and RATNER L. (1989) *Proc Natl Acad Sci USA* 86: 1128-1132.

PEPER RJ, TINA WZ, & MICHELSON MM (1968). *J Lab Clin Med* 72: 842-846.

25

SANGER F, NICKLEN S AND COULSON AR (1977). *Proc Natl Acad Sci USA* 4: 5463-67.

SHEPPARD HW, LANG W, ASCHER MS, VITTINGHOF E and WINKELSTEIN W.  
30 (1993) *AIDS* 7: 1159-1166.

SHIODA T, LEVY JA, CHENG-MAYER C (1992) *Proc Natl Acad Sci USA* 89: 9434-9438.

- 5 SHUGARS DC, SMITH MS, GLUECK DH, NANTERMET PV, SEILLIER-MOISEWITSCH F AND SWANSTROM R (1993). *J Virol* 67: 4639-4650.

SMITH J, AZAD AA, and DEACON NJ (1992) *J Gen Virol* 73: 1825-1828.

- 10 SULLIVAN N, THALI M, FURMAN C, HO DD, SODROSKI J (1993) *J Virol* 67: 3674-3679.

TEEUWSEN VJP, SIEBELINK KHJ, DE WOLF F, GOUDSMIT J, UYTDEHAAG FGCM, OSTERHAUS ADME (1990) *AIDS* 4: 77-81.

15.

TERSMETTE M, DE GOEDE REY, BERT JM, *et al* (1988) *J Virol* 62: 2026-2032.

TERWILLIGER E, SODROSKI JG, ROSEN CA, and HASELTINE WA, (1986) *J Virol* 60: 754-760.

20

YAMANOTO K MORI S, OKAMOTO T, SHINOTOHNO K AND KYOGOKU Y (1991). *Nucl Acids Res* 22: 6107-6112.

YANISCH-PERRON C, VIEIRA J, and MESSING J *Gene* 13: 103-119.



- 75 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i)(a) APPLICANT: (Other than US)

THE MACFARLANE BURNET CENTRE FOR MEDICAL RESEARCH LIMITED and  
AUSTRALIAN RED CROSS SOCIETY (NSW DIVISION)

## (b) INVENTORS/APPLICANTS (US only):

DEACON, N; LEARMONT, J; McPHEE, D; CROWE, S and COOPER, D.

## (ii) TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1

## (iii) NUMBER OF SEQUENCES: 800

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE

(B) STREET: 1 LITTLE COLLINS STREET

(C) CITY: MELBOURNE

(D) STATE: VICTORIA

(E) COUNTRY: AUSTRALIA

(F) ZIP: 3000

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT INTERNATIONAL

(B) FILING DATE: 14-FEB-1995

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PM3864 (AU)

(B) FILING DATE: 14-FEB-1994

(A) APPLICATION NUMBER: PM4002 (AU)

(B) FILING DATE: 21-FEB-1994

(A) APPLICATION NUMBER: PN0284 (AU)

(B) FILING DATE: 23-DEC-1994

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES, Dr E JOHN L

(C) REFERENCE/DOCKET NUMBER: EJH/EK

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 254 2777

(B) TELEFAX: +61 3 254 2770

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9709 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

TGGAAGGGCT AATTTGGTCC CAAAAAGAC AAGAGATCCT TGATCTGTGG ATCTACCACA    60
CACAGGCTA CTTCCTGAT TGGCAGAACT ACACACCAGG GCCAGGGATC AGATATCCAC    120
TGACCTTTGG ATGGTGCTTC AAGTTAGTAC CAGTTGAACC AGAGCAAGTA GAAGAGGCCA    180
AATAAGGAGA GAAGAACAGC TTGTTACACC CTATGAGCCA GCATGGGATG GAGGACCCGG    240
AGGGAGAAGT ATTAGTGTGG AAGTTTGACA GCCTCCTAGC ATTTGCTCAC ATGGCCCGAG    300
AGCTGCATCC GGAGTACTAC AAAGACTGCT GACATGAGC TTTCTACAAG GGACTTTCCG    360
CTGGGGACTT TCCAGGGAGG TGTGGCCTGG GCGGGACTGG GGAGTGGCGA GCCCTCAGAT    420
GCTACATATA AGCAGCTGCT TTTTGCCGTG ACTGGGTCTC TCTGGTTAGA CCAGATCTGA    480
GCCTGGGAGC TCTCTGGCTA ACTAGGGAAC CCAGTCTTA AGCCTCAATA AAGCTTGCCCT    540
TGAGTGCTCA AAGTAGTGTG TGCCCGTCTG TTGTGTGACT CTGGTAACTA GAGATCCCTC    600
AGACCCCTTT AGTCAGTGTG GAAAATCTCT AGCAGTGGCG CCCGAACAGG GACTTGAAAG    660
CGAAAGTAAA GCCAGAGGAG ATCTCTCGAC GCAGGACTCG GCTTGCTGAA GCGCGCACGG    720
CAAGAGGCGA GGGGCGGCGA CTGGTGAGTA CGCCAAAAT TTGACTAGC GGAGGCTAGA    780
AGGAGAGAGA TGGGTGCGAG AGCGTCGGTA TTAAGCGGGG GAGAATTAGA TAAATGGGAA    840
AAAATTGGGT TAAGGCCAGG GGGAAAGAAA CAATATAAAC TAAACATAT AGTATGGGCA    900
AGCAGGGAGC TAGAACGATT CGCAGTTAAT CCTGGCCTTT TAGAGACATC AGAAGGCTGT    960
AGACAAATAC TGGGACAGCT ACAACCATCC CTTGAGACAG GATCAGAAGA ACTTAGATCA   1020
TTATATAATA CAATAGCAGT CCTCTATTGT GTGCATCAAA GGATAGATGT AAAAGACACC   1080
AAGGAAGCCT TAGATAAGAT AGAGGAAGAG CAAAACAAA GTAAGAAAAA GGCACAGCAA   1140
GCAGCAGCTG ACACAGGAAA CAACAGCCAG GTCAGCCAAA ATTACCTTAT AGTCAGAAC   1200
CTCCAGGGGC AAATGGTACA TCAGGCCATA TCACCTAGAA CTTTAAATGC ATGGGTAAAA   1260
GTAGTAGAAG AGAAGGCTTT CAGCCAGAA GTAATACCCA TGTTTTGAGC ATTATCAGAA   1320
GGAGCCACCC CACAAGATTT AAATACCATG CTAACACAG TGGGGGGACA TCAAGCAGCC   1380
ATGCAATGT TAAAGAGAC CATCAATGAG GAAGCTGCAG AATGGGATAG ATTGCATCCA   1440
GTGCATGCGG GGCCTATTGC ACCAGGCCAG ATGAGAGAAC CAAGGGGAAG TGACATAGCA   1500
GGAACATACT GTACCCCTCA GGAACAAATA GGATGGATGA CACATAATCC ACCTATCCCA   1560
GTAGGAGAAA TCTATAAAG ATGGATAATC CTGGGATTAA ATAAAATAGT AAGAATGTAT   1620
AGCCCTACCA GCATTCTGGA CATAGACAA GGACCAAGG AACCCCTTAG AGACTATGTA   1680
GACCGATTCT ATAAACTCT AAGAGCCGAG CAAGCTTCAC AAGAGGTAAA AAATTGGATG   1740
ACAGAAACCT TGTTGGTCCA AAATGCGAAC CCAGATTGTA AGACTATTTT AAAAGCATTG   1800
GGACCAAGGAG CGACACTAGA AGAATGATG ACAGCATGTC AGGGAGTGGG GGGACCCGGC   1860
CATAAAGCAA GAGTTTGGC TGAAGCAATG AGCCAAGTAA CAAATCCAGC TACCATAATG   1920
ATACAGAAAG GCAATTTTAG GAACCAAGA AAGACTGTTA AGTGTTCCTA TTGTGGCAAA   1980
GAAGGGCACA TAGCCAAAA TTGCAGGGCC CCTAGGAAAA AGGGCTGTTG GAAATGTGGA   2040
AAGGAAGGAC ACCAAATGAA AGATTGTACT GAGAGACAGG CTAATTTTTT AGGGAAGATC   2100
TGGCCTTCCC ACAAGGAAG GCCAGGGAAT TTTCTTCAGA GCAGACCAGA GCCAACAGCC   2160
CCACCGAAG AGAGCTTCAG GTTTGGGGAA GAGACAACAA CTCCTCTCA GAAGCAGGAG   2220
CGATAGACA AGGAACGTGA TCCTTAGCT TCCCTCAGAT CACTCTTTGG CAGCGACCCC   2280

```

TCGTCACAAT AAAGATAGGG GGGCAATTAA AGGAAGCTCT ATTAGATACA GGAGCAGATG	2340
ATACAGTATT AGAAGAAATG AATTGGCCAG GAAGATGGAA ACCAAAAATG ATAGGGGGAA	2400
TTGGAGGTTT TATCAAAGTA GGACAGTATG ATCAGATACT CATAGAAATC TGCGGACATA	2460
AAGCTATAGG TACAGTATTA GTAGGACCTA CACCTGTCAA CATAATTGGA AGAAATCTGT	2520
TGACTCAGAT TGGCTGCACT TTAAATTTTC CCATTAGTCC TATTGAGACT GTACCAAGTAA	2580
AATTAAAGCC AGGAATGGAT GGCCCAAAAG TTAACAATG GCCATTGACA GAAGAAAAAA	2640
TAAAGCATT AGTAGAAATT TGTACAGAAA TGGAAAAGGA AGGAAAAATT TCAAAAATTG	2700
GGCCTGAAAA TCCATACAAT ACTCCAGTAT TTGCCATAAA GAAAAAGAC AGTACTAAAT	2760
GGAGAAAATT AGTAGATTTC AGAGAACTTA ATAAGAGAAC TCAAGATTTC TGGGAAGTTC	2820
AATTAGGAAT ACCACATCCT GCAGGGTTAA AACAGAAAAA ATCAGTAACA GTACTGGATG	2880
TGGGCGATGC ATATTTTTC GTTCCCTTAG ATAAAGACTT CAGGAAGTAT ACTGCAITTA	2940
CCATACCTAG TATAACAAT GAGACACCAG GGATTAGATA TCAGTACAAT GTGCTTCCAC	3000
AGGGATGGAA AGGATCACC GCAATATTCC AGTGTAGCAT GACAAAAATC TTAGAGCCTT	3060
TTAGAAAACA AAATCCAGAC ATAGTCATCT ATCAATACAT GGATGATTG TATGTAGGAT	3120
CTGACTTAGA AATAGGGCAG CATAGAACAA AAATAGAGGA ACTGAGACAA CATCTGTGA	3180
GGTGGGGATT TACCACACCA GACAAAAAAC ATCAGAAAGA ACCTCCATTC CTTTGGATGG	3240
GTTATGAAT CCATCCTGAT AAATGGACAG TACAGCCTAT AGTGTGCCA GAAAAGGACA	3300
GCTGGACTGT CAATGACATA CAGAAATTAG TGGGAAAAAT GAATTGGGCA AGTCAGATT	3360
ATGCAGGGAT TAAAGTAAGG CAATTATGTA AACTTCTTAG GGGAAACAAA GCACTAACAG	3420
AAGTAGTACC ACTAACAGAA GAAGCAGAGC TAGAACTGGC AGAAAACAGG GAGATTCTAA	3480
AGAACCCTGT ACATGGAGTG TATTATGACC CATCAAAAGA CTTAATAGCA GAAATACAGA	3540
AGCAGGGGCA AGGCCAATGG ACATATCAAA TTTATCAAGA GCCATTAAAA AATCTGAAAA	3600
CAGSAAATA TGCAAGATG AAGGGTGCC ACCTAATGA TGTGAACAA TTAACAGAGG	3660
CAGTACAAA AATAGCCACA GAAAGCATAG TAATATGGGG AAAGACTCCT AAATTTAAAT	3720
TACCATACA AAAGGAAACA TGGGAAGCAT GGTGGACAGA GTATTGGCAA GCCACCTGGA	3780
TTCTGAGTG GGAGTTTGT AATACCCCTC CCTTAGTGAA GTTATGGTAC CAGTTAGAGA	3840
AAGAACCCT AATAGGAGCA GAACTTTCT ATGTAGATGG GGCAGCCAAT AGGGAAGCTA	3900
AATTAGGAAA AGCAGGATAT GTAACCTGACA GAGGAAGACA AAAAGTTGTC CCCTAACGG	3960
ACACAACAAA TCAGAAGACT GAGTTACAAG CAATTCATCT AGCTTTGCAG GATTCGGGAT	4020
TAGAAGTAAA CATAGTGACA GACTCACAAT ATGCATTGGG AATCATTCAA GCACAACCAG	4080
ATAAGAGTGA ATCAGAGTTA GTCAGTCAA TAATAGAGCA GTTAATAAAA AAGGAAAAAG	4140
TCTACCTGGC ATGGGTACCA GCACACAAAG GAATTGGAGG AAATGAACAA GTAGATGGGT	4200
TGGTCAGTGC TGGAAACAG AAGTACTAT TTTTAGATGG AATAGATAAG GCCCAAGAAG	4260
AACATGAGAA ATATCACAGT AATTGGAGAG CAATGGCTAG TGATTTTAACT ACCACCTG	4320
TAGTAGCAAA AGAAATAGTA GCCAGCTGTG ATAAATGTCA GCTAAAAGGG GAAGCCATGC	4380
ATGGACAAGT AGACTGTAGC CCAGGAATAT GGCAGCTAGA TTGTACACAT TTAGAAGGAA	4440
AAGTTATCTT GGTAGCAGTT CATGTAGCCA GTGGATATAT AGAAGCAGAA GTAATTCCAG	4500
CAGAGACAGG GCAAGAAACA GCATACTTCC TCTTAAAAAT AGCAGGAAGA TGGCCAGTAA	4560
AAACAGTACA TACAGACAAT GGCAGCAATT TCACCACTAC TACAGTTAAG GCGCCTGTT	4620
GGTGGGCGGG GATCAAGCAG GAATTGGCA TTCCCTACAA TCCCAAAAGT CAAGGAGTAA	4680
TAGAATCTAT GAATAAGAA TTAAGAAAA TTATAGGACA GGTAAGAGAT CAGGCTGAAC	4740
ATCTTAAGAC AGCAGTACAA ATGGCAGTAT TCATCCCAA TTTTAAAAAG AAAGGGGGGA	4800
TTGGGGGGTA CAGTGCAGGG GAAAGAATAG TAGACATAAT AGCAACAGAC ATACAACTA	4860
AAGAATTACA AAAACAATT ACAAATTC AAAATTTTCG GGTATTATAC AGGACAGCA	4920
GAGATCCAGT TTGGAAGGA CCAGCAAGC TCCTCTGGAA AGGTGAAGGG GCAGTAGTAA	4980
TACAAGATAA TAGTGACATA AAGTAGTGC CAAGAAGAAA AGCAAGATC ATCAGGGATT	5040
ATGGAACA GATGGCAGT GATGATTGTG TGGCAAGTAG ACAGGATGAG GATTACACA	5100

TGGAAAAGAT TAGTAAACA CCATATGTAT ATTTCAAGGA AAGCTAAGGA CTGGTTTTAT	5160
AGACATCACT ATGAAAGTAC TAATCCAAA ATAAGTTCAG AAGTACACAT CCCACTAGGG	5220
GATGCTAAAT TAGTAATAAC AACATATTGG GGTCTGCATA CAGGAGAAAG AGACTGGCAT	5280
TTGGGTCAGG GAGTCTCCAT AGAATGGAGG AAAAAGAGAT ATAGCACACA AGTAGACCCT	5340
GACCTAGCAG ACCAACTAAT TCATCTGCAC TATTTTGATT GTTTTTCAGA ATCTGCTATA	5400
AGAAATACCA TATTAGGACG TATAGTTAGT CCTAGGTGTG AATATCAAGC AGGACATAAC	5460
AAGGTAGGAT CTCTACAGTA CTGGGCACTA GCAGCATTAA TAAAACCAA ACAGATAAAG	5520
CCACCTTTGC CTAGTGTAG GAAACTGACA GAGGACAGAT GGAACAAGCC CCAGAAGACC	5580
AAGGGCCACA GAGGGAGCCA TACAATGAAT GGACACTAGA GCTTTTAGAG GAACCTAAGA	5640
GTGAAGCTGT TAGACATTTT CCTAGGATAT GGCTCCATAA CTTAGGACAA CATATCTATG	5700
AAACTTACGG GGATACTTGG GCAGGAGTGG AAGCCATAAT AAGAATTCTG CAACAACCTGC	5760
TGTTTTCCA TTTCAGAAAT GGGTGTGAC ATAGCAGAAT AGGCGTACT CGACAGAGGA	5820
GAGCAAGAAA TGGAGCCAGT AGATCCTAGA CTAGAGCCCT GGAAGCATCC AGGAAGTCAG	5880
CCTAAAAGT CTGTACCAA TTGCTATTGT AAAAAGTGT GCTTTCATTG CCAAGTTTGT	5940
TTCATGACAA AAGCCTTAGG CATCTCCTAT GGCAGGAAGA AGCGGAGACA GCGACGAAGA	6000
GCTCATCAGA ACASTCAGAC TCATCAAGCT TCTCTATCAA AGCAGTAAGT AGTACATGTA	6060
ATGCAACCTA TAATAGTAGC AATAGTAGCA TTAGTAGTAG CAATAATAAT AGCAATAGTT	6120
GTGTGGTCCA TAGTAATCAT AGAATATAGG AAAATATTAA GACAAAGAAA AATAGACAGG	6180
TTAATTGATA GACTAATAGA AAGAGCAGAA GACAGTGGCA ATGAGAGTGA AGGAGAAGTA	6240
TCAGCACTTG TGGAGATGGG GGTGGAAATG GGGCACCATG CTCCTTGGGA TATTGATGAT	6300
CTGTAGTGCT ACAGAAAAAT TGTGGGTCAC AGTCTATTAT GGGGTACCTG TGTGGAAGGA	6360
AGCAACCACC ACTCTATTTT GTGCATCAGA TGCTAAAGCA TATGATACAG AGGTACATAA	6420
TGTTTGGGCC ACACATGCCT GTGTACCCAC AGACCCCAAC CCACAAGAAG TAGTATTGGT	6480
AAATGTGACA GAAATTTTTA ACATGTGGAA AAATGACATG GTAGAACAGA TGCATGAGGA	6540
TATAATCAGT TTATGGGATC AAAGCCTAAA GCCATGTGTA AAATTAACCC CACTCTGTGT	6600
TAGTTTAAAG TGCATGATT TGAAGATGA TACTAATACC AATAGTAGTA GCGGAGAAAT	6660
GATAATGGAG AAAGGAGAGA TAAAAAAGT CTCTTTCAAT ATCAGCACA GCATAAGAGA	6720
TAAAGTGACG AAAGAATATG CATTCTTTTA TAAACTTGAT ATAGTACCAA TAGATAATAC	6780
CAGCTATAGG TTGATAAGTT GTAACACCTC AGTCATTACA CAGGCTGTC CAAAGGTATC	6840
CTTTGAGCCA ATTCCATAC ATTATTGTGC CCGGCTGGT TTTGGGATTC TAAATGTAA	6900
TAATAAGACG TTCAATGGAA CAGGACCATG TACAAATGTC AGCAGATAC AATGTACACA	6960
TGGAATCAGG CCAGTAGTAT CAACTCACT GCTGTAAAT GGCAGTCTAG CAGAAGAAGA	7020
TGTAGTAATT AGATCTGCCA ATTTACAGA CAATGCTAAA ACCATAATAG TACAGCTGAA	7080
CACATCTGTA GAAATTAATT GTACAGACC CAACAACAT ACAAGAAAA GTATCCGTAT	7140
CCAGAGGGGA CCAGGAGAG CATTGTATC AATAGGAAAA ATAGGAAATA TGAGACAAGC	7200
ACATTGTAAC ATTAGTAGAG CAAATGGAA TGCCACTTTA AAACAGATAG CTAGCAAAT	7260
AAGAGACAA TTTGGAAATA ATAAAACAT AATCTTAAAG CAATCCTCAG GAGGGGACCC	7320
AGAAATTGTA ACGCACAGTT TTAATTGTGG AGGGGAATTT TTCTACTGTA ATTCAACACA	7380
ACTGTTTAAT AGTACTTGGT TTAATAGTAC TTGGAGTACT GAAGGGTCAA ATAACACTGA	7440
AGGAAGTGAC ACAATCACAC TCCCATGCAG AATAAACAA TTTATAACA TGTGGCAGGA	7500
AGTAGGAAAA GCAATGTATG CCCCTCCCAT CAGTGGACAA ATTAGATGTT CATCAAATAT	7560
TACTGGGCTG CTATTAACAA GAGATGGTGG TAATAACAAC AATGGGTCG AGATCTTCAG	7620
ACCTGGAGGA GCGGATATGA GGCACAATTG GAGAAGTGAA TTATATAAAT ATAAAGTAGT	7680
AAAAATTGAA CCATTAGGAG TAGCACCCAC CAAGGCCAAG AGAAGAGTGG TGCAGAGAGA	7740
AAAAAGAGCA GTGGGAATAG GAGCTTTGTT CCTTGGGTTT TTGGGAGCAG CAGGAAGCAC	7800
TATGGGCTGC ACGTCAATGA CGCTGACGGT ACAGGCCAGA CAATTATTGT CTGATATAGT	7860
GCAGCAGCAG AACAATTTGC TGAGGGCTAT TGAGGCGCAA CAGCATCTGT TGCAACTCAC	7920

AGTCTGGGGC ATCAAACAGC TCCAGGCAAG AATCCTGGCT GTGAAAAGAT ACCTAAAGGA	7980
TCAAACAGCTC CTGGGGATTT GGGGTGCTC TGGAAAAC TC ATTTGCACCA CTGCTGTGCC	8040
TTGGAATGCT AGTTGGAGTA ATAAATCTCT GGAACAGATT TGGAAATA CA TGACCTGGAT	8100
GGAGTGGGAC AGAGAAATTA ACAATTACAC AAGCTTAATA CACTCCTTAA TTGAAGAATC	8160
GCAAAACCAG CAAGAAAAGA ATGAACAAGA ATTATTGGAA TTAGATAAAT GGGCAAGTTT	8220
GTGGAATTGG TTAAACATAA CAAATTGGCT GTGGTATATA AAATTATTCA TAATGATAGT	8280
AGGAGGCTTG GTAGSTTTAA GAATAGTTTT TGCTGTACTT TCTATAGTGA ATAGAGTTAG	8340
GCAGGGATAT TCACCAATTAT CGTTTCAGAC CCACCTCCCA ATCCCGAGGG GACCCGACAG	8400
GCCCGAAGGA ATAGAAGAAG AAGGTGGAGA GAGAGACAGA GACAGATCCA TTCGATTAGT	8460
GAACGGATCC TTAGCACTTA TCTGGGACGA TCTGCGGAGC CTGTGCCTCT TCAGCTACCA	8520
CCGCTTGAGA GACTTACTCT TGATTGTAAC GAGGATTGTG GAACCTCTGG GACCGAGGGG	8580
GTGGGAAGCC CTCAAATATT GGTGGAATCT CTTACAGTAT TGGAGTCAGG AACTAAGAA	8640
TAGTGCTGTT AACTTGCTCA ATGCCACAGC CATAGCAGTA GCTGAGGGGA CAGATAGGGT	8700
TATAGAGTA TTACAAGCAG CTTATAGAGC TATTGCGCAC ATACCTAGAA GAATAAGACA	8760
GGGCTTGGA AGGATTTTGC TATAAGATGG GTGGCAAGTG GTCAAAAAGT AGTGTGATTG	8820
GATGGCCTGC TGTAAAGGAA AGAATGAGAC GAGCTGAGCC AGCAGCAGAT GGGGTGGGAG	8880
CAGTATCTCG AGACCTAGAA AAACATGGAG CAATCACAG TAGCAATACA GCAGCTAACA	8940
ATGCTGCTTG TGCTGGCTA GAAGCACAAG AGGAGGAAGA GGTGGGTTTT CCAGTCACAC	9000
CTCAGGTACC TTTAAGACCA ATGACTTACA AGGCAGCTGT AGATCTTAGC CACTTTTTAA	9060
AAGAAAAGGG GGGACTGGAA GGGCTAATTC ACTCCCAAAG AAGACAAGAT ATCCTTGATC	9120
TGTGGATCTA CCACACACAA GGCTACTTCC CTGATTGGCA GAACTACACA CCAGGGCCAG	9180
GGGTCAGATA TCCACTGACC TTGGATGGT GCTACAAGCT AGTACCAGTT GAGCCAGATA	9240
AGGTAGAAGA GGCCAATAA GGAGAGAACA CCAGCTTGT ACACCTGTG AGCCTGCATG	9300
GAATGGATGA CCGTAGAGA GAAGTGTAG AGTGGAGGTT TGACAGCCGC CTAGCATTTT	9360
ATCACGTGGC CCGAGAGCTG CATCCGAGT ACTTCAAGAA CTGCTGACAT CGAGCTTGCT	9420
ACAAGGGACT TTCGCTGGG GACTTTCCAG GGAGGCGTGG CCTGGGCGGG ACTGGGGAGT	9480
GGCGAGCCCT CAGATGCTGC ATATAAGCAG CTGCTTTTTG CCTGTACTGG GTCTCTCTGG	9540
TTAGACCAGA TCTGAGCCTG GGAGCTCTCT GGCTAACTAG GGAACCCACT GCTTAAGCCT	9600
CAATAAGCT TGCCCTTGAGT GCTTCAAGTA GTGTGTGCC GTCTGTGTG TGACTCTGGT	9660
AACTAGAGAT CCCTCAGACC CTTTATGTC GTGTGAAAA TCTCTAGCA	9709

- 80 -

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGGGTGGCA

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGGGTGGCAA

## (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGTGGCAAG

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGTGGCAAGT

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTGGCAAGTG

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGGCAAGTGG

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGCAAGTGGT

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCAAGTGGTC

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

- 81 -

CAAGTGGTCA

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AAGTGGTCAA

## (2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGTGGTCAAA

## (2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GTGGTCAAAA

## (2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGGTCAAAAA

## (2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGTCAAAAAG

## (2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTCAAAAAGT

## (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TCAAAAAGTA

## (2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CAAAAAGTAG

## (2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AAAAAGTAGT

- 82 -

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AAAAGTAGTG

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AAAGTAGTGT

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AAGTAGTGTG

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AGTAGTGTGA

## (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTAGTGTGAT

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TAGTGTGATT

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AGTGTGATTG

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GTGTGATTGG

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TGTGATTGGA

## (2) INFORMATION FOR SEQ ID NO:29:



- 83 -

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GTGATTGGAT

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TGATTGGATG

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GATTGGATGG

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATTGGATGGC

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TTGGATGGCC

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TGGATGGCCT

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GGATGGCCTG

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GATGGCCTGC

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

ATGGCCTGCT

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:

- 84 -

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TGGCCTGCTG

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GGCCTGCTGT

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GCCTGCTGTA

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CCTGCTGTAA

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CTGCTGTAAG

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

TGCTGTAAGG

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GCTGTAAGGG

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CTGTAAGGGA

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TGTAAGGGAA

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid

- 85 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GTAAGGGAAA

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TAAGGGAAAG

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AAGGGAAAGA

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGGGAAAGAA

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GGGAAAGAAT

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GGAAAGAATG

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GAAAGAATGA

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

AAAGAATGAG

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

AAGAATGAGA

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 86 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

AGAATGAGAC

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GAATGAGACG

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

AATGAGACGA

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

ATGAGACGAG

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TGAGACGAGC

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GAGACGAGCT

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

AGACGAGCTG

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

GACGAGCTGA

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

ACGAGCTGAG

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 87 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

CGAGCTGAGC

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GAGCTGAGCC

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

AGCTGAGCCA

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GCTGAGCCAG

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

CTGAGCCAGC

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

TGAGCCAGCA

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

GAGCCAGCAG

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

AGCCAGCAGC

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

GCCAGCAGCA

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

- 88 -

CCAGCAGCAG

## (2) INFORMATION FOR SEQ ID NO:75:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

CAGCAGCAGA

## (2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

AGCAGCAGAT

## (2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

GCAGCAGATG

## (2) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

CAGCAGATGG

## (2) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

AGCAGATGGG

## (2) INFORMATION FOR SEQ ID NO:80:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

GCAGATGGGG

## (2) INFORMATION FOR SEQ ID NO:81:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

CAGATGGGGT

## (2) INFORMATION FOR SEQ ID NO:82:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

AGATGGGGTG

## (2) INFORMATION FOR SEQ ID NO:83:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

GATGGGGTGG

- 89 -

## (2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

ATGGGGTGGG

## (2) INFORMATION FOR SEQ ID NO:85:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

TGGGGTGGGA

## (2) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

GGGGTGGGAG

## (2) INFORMATION FOR SEQ ID NO:87:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

GGGTGGGAGC

## (2) INFORMATION FOR SEQ ID NO:88:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

GGTGGGAGCA

## (2) INFORMATION FOR SEQ ID NO:89:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

GTGGGAGCAG

## (2) INFORMATION FOR SEQ ID NO:90:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

TGGGAGCAGT

## (2) INFORMATION FOR SEQ ID NO:91:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

GGGAGCAGTA

## (2) INFORMATION FOR SEQ ID NO:92:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

GGAGCAGTAT

## (2) INFORMATION FOR SEQ ID NO:93:

- 90 -

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

GAGCAGTATC

(2) INFORMATION FOR SEQ ID NO:94:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

AGCAGTATCT

(2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

GCAGTATCTC

(2) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

CAGTATCTCG

(2) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

AGTATCTCGA

(2) INFORMATION FOR SEQ ID NO:98:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

GTATCTCGAG

(2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

TATCTCGAGA

(2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

ATCTCGAGAC

(2) INFORMATION FOR SEQ ID NO:101:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

TCTCGAGACC

(2) INFORMATION FOR SEQ ID NO:102:

- (i) SEQUENCE CHARACTERISTICS:



- 91 -

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

CTCGAGACCT

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

TCGAGACCTA

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

CGAGACCTAG

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

GAGACCTAGA

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

AGACCTAGAA

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

GACCTAGAAA

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

ACCTAGAAAA

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

CCTAGAAAAA

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

CTAGAAAAAC

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid

- 92 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

TAGAAAAACA

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

AGAAAAACAT

(2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

GAAAAACATG

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

AAAAACATGG

(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

AAAACATGGA

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

AAACATGGAG

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

AACATGGAGC

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

ACATGGAGCA

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

CATGGAGCAA

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 93 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:  
ATGGAGCAAT

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:  
TGGAGCAATC

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:  
GGAGCAATCA

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:  
GAGCAATCAC

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:  
AGCAATCACA

(2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:  
GCAATCACAA

(2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:  
CAATCACAAG

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:  
AATCACAAGT

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:  
ATCACAAGTA

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 94 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:  
TCACAAGTAG

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:  
CACAAGTAGC

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:  
ACAAGTAGCA

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:  
CAAGTAGCAA

(2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:  
AAGTAGCAAT

(2) INFORMATION FOR SEQ ID NO:134:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:  
AGTAGCAATA

(2) INFORMATION FOR SEQ ID NO:135:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:  
GTAGCAATAC

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:  
TAGCAATACA

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:  
AGCAATACAG

(2) INFORMATION FOR SEQ ID NO:138:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

- 95 -

GCAATACAGC

## (2) INFORMATION FOR SEQ ID NO:139:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

CAATACAGCA

## (2) INFORMATION FOR SEQ ID NO:140:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

AATACAGCAG

## (2) INFORMATION FOR SEQ ID NO:141:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

ATACAGCAGC

## (2) INFORMATION FOR SEQ ID NO:142:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

TACAGCAGCT

## (2) INFORMATION FOR SEQ ID NO:143:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

ACAGCAGCTA

## (2) INFORMATION FOR SEQ ID NO:144:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

CAGCAGCTAA

## (2) INFORMATION FOR SEQ ID NO:145:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

AGCAGCTAAC

## (2) INFORMATION FOR SEQ ID NO:146:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

GCAGCTAACA

## (2) INFORMATION FOR SEQ ID NO:147:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

CAGCTAACAA

- 96 -

## (2) INFORMATION FOR SEQ ID NO:148:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

AGCTAACAAAT

## (2) INFORMATION FOR SEQ ID NO:149:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

GCTAACAAATG

## (2) INFORMATION FOR SEQ ID NO:150:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

CTAACAAATGC

## (2) INFORMATION FOR SEQ ID NO:151:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

TAACAATGCT

## (2) INFORMATION FOR SEQ ID NO:152:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

AACAAATGCTG

## (2) INFORMATION FOR SEQ ID NO:153:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

ACAATGCTGC

## (2) INFORMATION FOR SEQ ID NO:154:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

CAATGCTGCT

## (2) INFORMATION FOR SEQ ID NO:155:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

AATGCTGCTT

## (2) INFORMATION FOR SEQ ID NO:156:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

ATGCTGCTTG

## (2) INFORMATION FOR SEQ ID NO:157:

- 97 -

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

TGCTGCTTGT

(2) INFORMATION FOR SEQ ID NO:158:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

GCTGCTTGTG

(2) INFORMATION FOR SEQ ID NO:159:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

CTGCTTGTGC

(2) INFORMATION FOR SEQ ID NO:160:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

TGCTTGTGCC

(2) INFORMATION FOR SEQ ID NO:161:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

GCTTGTGCCT

(2) INFORMATION FOR SEQ ID NO:162:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

CTTGTGCCTG

(2) INFORMATION FOR SEQ ID NO:163:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

TTGTGCCTGG

(2) INFORMATION FOR SEQ ID NO:164:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

TGTGCCTGGC

(2) INFORMATION FOR SEQ ID NO:165:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

GTGCCTGGCT

(2) INFORMATION FOR SEQ ID NO:166:

- (i) SEQUENCE CHARACTERISTICS:

- 98 -

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

TGCCTGGCTA

(2) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

GCCTGGCTAG

(2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

CCTGGCTAGA

(2) INFORMATION FOR SEQ ID NO:169:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

CTGGCTAGAA

(2) INFORMATION FOR SEQ ID NO:170:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

TGGCTAGAAG

(2) INFORMATION FOR SEQ ID NO:171:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

GGCTAGAAGC

(2) INFORMATION FOR SEQ ID NO:172:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

GCTAGAAGCA

(2) INFORMATION FOR SEQ ID NO:173:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

CTAGAAGCAC

(2) INFORMATION FOR SEQ ID NO:174:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

TAGAAGCACA

(2) INFORMATION FOR SEQ ID NO:175:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid



- 99 -

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

AGAAGCACAA

(2) INFORMATION FOR SEQ ID NO:176:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

GAAGCACAAG

(2) INFORMATION FOR SEQ ID NO:177:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

AAGCACAAGA

(2) INFORMATION FOR SEQ ID NO:178:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

AGCACAAGAG

(2) INFORMATION FOR SEQ ID NO:179:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

GCACAAGAGG

(2) INFORMATION FOR SEQ ID NO:180:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

CACAAGAGGA

(2) INFORMATION FOR SEQ ID NO:181:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

ACAAGAGGAG

(2) INFORMATION FOR SEQ ID NO:182:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

CAAGAGGAGG

(2) INFORMATION FOR SEQ ID NO:183:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

AAGAGGAGGA

(2) INFORMATION FOR SEQ ID NO:184:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- 100 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

AGAGGAGGAA

(2) INFORMATION FOR SEQ ID NO:185:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

GAGGAGGAAG

(2) INFORMATION FOR SEQ ID NO:186:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

AGGAGGAAGA

(2) INFORMATION FOR SEQ ID NO:187:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

GGAGGAAGAG

(2) INFORMATION FOR SEQ ID NO:188:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

GAGGAAGAGG

(2) INFORMATION FOR SEQ ID NO:189:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

AGGAAGAGGT

(2) INFORMATION FOR SEQ ID NO:190:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

GGAAGAGGTG

(2) INFORMATION FOR SEQ ID NO:191:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

GAAGAGGTGG

(2) INFORMATION FOR SEQ ID NO:192:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

AAGAGGTGGG

(2) INFORMATION FOR SEQ ID NO:193:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 101 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:  
AGAGGTGGGT

(2) INFORMATION FOR SEQ ID NO:194:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:  
GAGGTGGGTT

(2) INFORMATION FOR SEQ ID NO:195:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:  
AGGTGGGTTT

(2) INFORMATION FOR SEQ ID NO:196:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:  
GGTGGGTTTT

(2) INFORMATION FOR SEQ ID NO:197:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:  
GTGGGTTTTTC

(2) INFORMATION FOR SEQ ID NO:198:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:  
TGGGTTTTTC

(2) INFORMATION FOR SEQ ID NO:199:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:  
GGGTTTTCCA

(2) INFORMATION FOR SEQ ID NO:200:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:200:  
GGTTTTCCAG

(2) INFORMATION FOR SEQ ID NO:201:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:  
GTTTTCCAGT

(2) INFORMATION FOR SEQ ID NO:202:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

- 102 -

TTTTCCAGTC

## (2) INFORMATION FOR SEQ ID NO:203:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

TTTCCAGTCA

## (2) INFORMATION FOR SEQ ID NO:204:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

TTCCAGTCAC

## (2) INFORMATION FOR SEQ ID NO:205:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:

TCCAGTCACA

## (2) INFORMATION FOR SEQ ID NO:206:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:206:

CCAGTCACAC

## (2) INFORMATION FOR SEQ ID NO:207:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:207:

CAGTCACACC

## (2) INFORMATION FOR SEQ ID NO:208:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:

AGTCACACCT

## (2) INFORMATION FOR SEQ ID NO:209:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

GTCACACCTC

## (2) INFORMATION FOR SEQ ID NO:210:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:210:

TCACACCTCA

## (2) INFORMATION FOR SEQ ID NO:211:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:211:

CACACCTCAG

- 103 -

## (2) INFORMATION FOR SEQ ID NO:212:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:

ACACCTCAGG

## (2) INFORMATION FOR SEQ ID NO:213:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:

CACCTCAGGT

## (2) INFORMATION FOR SEQ ID NO:214:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:

ACCTCAGGTA

## (2) INFORMATION FOR SEQ ID NO:215:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:215:

CCTCAGGTAC

## (2) INFORMATION FOR SEQ ID NO:216:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:216:

CTCAGGTACC

## (2) INFORMATION FOR SEQ ID NO:217:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:217:

TCAGGTACCT

## (2) INFORMATION FOR SEQ ID NO:218:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:218:

CAGGTACCTT

## (2) INFORMATION FOR SEQ ID NO:219:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:

AGGTACCTTT

## (2) INFORMATION FOR SEQ ID NO:220:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:

GGTACCTTTA

## (2) INFORMATION FOR SEQ ID NO:221:

- 104 -

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:

GTACCTTTAA

(2) INFORMATION FOR SEQ ID NO:222:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:

TACCTTTAAG

(2) INFORMATION FOR SEQ ID NO:223:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:

ACCTTTAAGA

(2) INFORMATION FOR SEQ ID NO:224:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

CCTTTAAGAC

(2) INFORMATION FOR SEQ ID NO:225:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

CTTTAAGACC

(2) INFORMATION FOR SEQ ID NO:226:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:226:

TTTAAGACCA

(2) INFORMATION FOR SEQ ID NO:227:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:227:

TTAAGACCAA

(2) INFORMATION FOR SEQ ID NO:228:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:228:

TAAGACCAAT

(2) INFORMATION FOR SEQ ID NO:229:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:229:

AAGACCAATG

(2) INFORMATION FOR SEQ ID NO:230:

(i) SEQUENCE CHARACTERISTICS:

- 105 -

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:230:

AGACCAATGA

(2) INFORMATION FOR SEQ ID NO:231:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:231:

GACCAATGAC

(2) INFORMATION FOR SEQ ID NO:232:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:232:

ACCAATGACT

(2) INFORMATION FOR SEQ ID NO:233:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:233:

CCAATGACTT

(2) INFORMATION FOR SEQ ID NO:234:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:234:

CAATGACTTA

(2) INFORMATION FOR SEQ ID NO:235:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:235:

AATGACTTAC

(2) INFORMATION FOR SEQ ID NO:236:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:236:

ATGACTTACA

(2) INFORMATION FOR SEQ ID NO:237:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:237:

TGACTTACAA

(2) INFORMATION FOR SEQ ID NO:238:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:238:

GACTTACAAG

(2) INFORMATION FOR SEQ ID NO:239:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid

- 106 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:239:

ACTTACAAGG

(2) INFORMATION FOR SEQ ID NO:240:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:240:

CTTACAAGGC

(2) INFORMATION FOR SEQ ID NO:241:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:241:

TTACAAGGCA

(2) INFORMATION FOR SEQ ID NO:242:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:242:

TACAAGGCAG

(2) INFORMATION FOR SEQ ID NO:243:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:243:

ACAAGGCAGC

(2) INFORMATION FOR SEQ ID NO:244:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:244:

CAAGGCAGCT

(2) INFORMATION FOR SEQ ID NO:245:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245:

AAGGCAGCTG

(2) INFORMATION FOR SEQ ID NO:246:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:246:

AGGCAGCTGT

(2) INFORMATION FOR SEQ ID NO:247:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:247:

GGCAGCTGTA

(2) INFORMATION FOR SEQ ID NO:248:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear



- 107 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:248:

GCAGCTGTAG

(2) INFORMATION FOR SEQ ID NO:249:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:249:

CAGCTGTAGA

(2) INFORMATION FOR SEQ ID NO:250:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:250:

AGCTGTAGAT

(2) INFORMATION FOR SEQ ID NO:251:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:251:

GCTGTAGATC

(2) INFORMATION FOR SEQ ID NO:252:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:252:

CTGTAGATCT

(2) INFORMATION FOR SEQ ID NO:253:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:253:

TG TAGATCTT

(2) INFORMATION FOR SEQ ID NO:254:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:254:

GTAGATCTTA

(2) INFORMATION FOR SEQ ID NO:255:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:255:

TAGATCTTAG

(2) INFORMATION FOR SEQ ID NO:256:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:

AGATCTTAGC

(2) INFORMATION FOR SEQ ID NO:257:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 108 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:257:  
GATCTTAGCC

(2) INFORMATION FOR SEQ ID NO:258:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:258:  
ATCTTAGCCA

(2) INFORMATION FOR SEQ ID NO:259:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:259:  
TCTTAGCCAC

(2) INFORMATION FOR SEQ ID NO:260:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:260:  
CTTAGCCACT

(2) INFORMATION FOR SEQ ID NO:261:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:261:  
TTAGCCACTT

(2) INFORMATION FOR SEQ ID NO:262:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:262:  
TAGCCACTTT

(2) INFORMATION FOR SEQ ID NO:263:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:263:  
AGCCACTTTT

(2) INFORMATION FOR SEQ ID NO:264:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:264:  
GCCACTTTTT

(2) INFORMATION FOR SEQ ID NO:265:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:265:  
CCACTTTTTA

(2) INFORMATION FOR SEQ ID NO:266:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:266:

- 109 -

CACTTTTAA

## (2) INFORMATION FOR SEQ ID NO:267:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:267:

ACTTTTAA

## (2) INFORMATION FOR SEQ ID NO:268:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:268:

CTTTTAA

## (2) INFORMATION FOR SEQ ID NO:269:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:269:

TTTTTAAAG

## (2) INFORMATION FOR SEQ ID NO:270:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:270:

TTTTTAAAGA

## (2) INFORMATION FOR SEQ ID NO:271:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:271:

TTTAAAAGAA

## (2) INFORMATION FOR SEQ ID NO:272:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:272:

TTAAAAGAAA

## (2) INFORMATION FOR SEQ ID NO:273:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:273:

TAAAAGAAAA

## (2) INFORMATION FOR SEQ ID NO:274:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:274:

AAAAGAAAAG

## (2) INFORMATION FOR SEQ ID NO:275:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:275:

AAAGAAAAGG

- 110 -

## (2) INFORMATION FOR SEQ ID NO:276:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:276:

AAGAAAAGGG

## (2) INFORMATION FOR SEQ ID NO:277:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:277:

AGAAAAGGGG

## (2) INFORMATION FOR SEQ ID NO:278:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:278:

GAAAAGGGGG

## (2) INFORMATION FOR SEQ ID NO:279:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:279:

AAAAGGGGGG

## (2) INFORMATION FOR SEQ ID NO:280:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:280:

AAAGGGGGGA

## (2) INFORMATION FOR SEQ ID NO:281:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:281:

AAGGGGGGAC

## (2) INFORMATION FOR SEQ ID NO:282:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:282:

AGGGGGGACT

## (2) INFORMATION FOR SEQ ID NO:283:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:283:

GGGGGGACTG

## (2) INFORMATION FOR SEQ ID NO:284:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:284:

GGGGGACTGG

## (2) INFORMATION FOR SEQ ID NO:285:

- 111 -

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:285:

GGGGACTGGA

(2) INFORMATION FOR SEQ ID NO:286:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:286:

GGGACTGGAA

(2) INFORMATION FOR SEQ ID NO:287:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:287:

GGACTGGAAG

(2) INFORMATION FOR SEQ ID NO:288:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:288:

GACTGGAAGG

(2) INFORMATION FOR SEQ ID NO:289:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:289:

ACTGGAAGGG

(2) INFORMATION FOR SEQ ID NO:290:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:290:

CTGGAAGGGC

(2) INFORMATION FOR SEQ ID NO:291:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:291:

TGGAAGGGCT

(2) INFORMATION FOR SEQ ID NO:292:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:292:

GGAAGGGCTA

(2) INFORMATION FOR SEQ ID NO:293:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:293:

GAAGGGCTAA

(2) INFORMATION FOR SEQ ID NO:294:

- (i) SEQUENCE CHARACTERISTICS:

- 112 -

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:294:

AAGGGCTAAT

(2) INFORMATION FOR SEQ ID NO:295:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:295:

AGGGCTAATT

(2) INFORMATION FOR SEQ ID NO:296:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:296:

GGGCTAATTC

(2) INFORMATION FOR SEQ ID NO:297:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:297:

GGCTAATTCA

(2) INFORMATION FOR SEQ ID NO:298:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:298:

GCTAATTCAC

(2) INFORMATION FOR SEQ ID NO:299:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:299:

CTAATTCAC

(2) INFORMATION FOR SEQ ID NO:300:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:300:

TAATTCAC

(2) INFORMATION FOR SEQ ID NO:301:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:301:

AATTCAC

(2) INFORMATION FOR SEQ ID NO:302:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:302:

ATTCAC

(2) INFORMATION FOR SEQ ID NO:303:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid

- 113 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:303:

TTCAC TCCCA

(2) INFORMATION FOR SEQ ID NO:304:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:304:

TCACTCCCAA

(2) INFORMATION FOR SEQ ID NO:305:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:305:

CACTCCCAAA

(2) INFORMATION FOR SEQ ID NO:306:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:306:

ACTCCCAAAG

(2) INFORMATION FOR SEQ ID NO:307:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:307:

CTCCCAAAGA

(2) INFORMATION FOR SEQ ID NO:308:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:308:

TCCCAAAGAA

(2) INFORMATION FOR SEQ ID NO:309:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:309:

CCCAAAGAAG

(2) INFORMATION FOR SEQ ID NO:310:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:310:

CCAAAGAAGA

(2) INFORMATION FOR SEQ ID NO:311:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:311:

CAAAGAAGAC

(2) INFORMATION FOR SEQ ID NO:312:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 114 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:312:

AAAGAAGACA

(2) INFORMATION FOR SEQ ID NO:313:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:313:

AAGAAGACAA

(2) INFORMATION FOR SEQ ID NO:314:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:314:

AGAAGACAAG

(2) INFORMATION FOR SEQ ID NO:315:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:315:

GAAGACAAGA

(2) INFORMATION FOR SEQ ID NO:316:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:316:

AAGACAAGAT

(2) INFORMATION FOR SEQ ID NO:317:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:317:

AGACAAGATA

(2) INFORMATION FOR SEQ ID NO:318:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:318:

GACAAGATAT

(2) INFORMATION FOR SEQ ID NO:319:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:319:

ACAAGATATC

(2) INFORMATION FOR SEQ ID NO:320:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:320:

CAAGATATCC

(2) INFORMATION FOR SEQ ID NO:321:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA



- 115 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:321:

AAGATATCCT

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(2) INFORMATION FOR SEQ ID NO:322:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:322:

AGATATCCTT

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:326:

ATCCTTGATC

(2) INFORMATION FOR SEQ ID NO:327:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:327:

TCCTTGATCT

(2) INFORMATION FOR SEQ ID NO:323:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:323:

GATATCCTTG

(2) INFORMATION FOR SEQ ID NO:328:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:328:

CCTTGATCTG

(2) INFORMATION FOR SEQ ID NO:324:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:324:

ATATCCTTGA

(2) INFORMATION FOR SEQ ID NO:329:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:329:

CTTGATCTGT

(2) INFORMATION FOR SEQ ID NO:325:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:325:

TATCCTTGAT

(2) INFORMATION FOR SEQ ID NO:330:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:330:

(2) INFORMATION FOR SEQ ID NO:326:

(i) SEQUENCE CHARACTERISTICS:

- 116 -

TTGATCTGTG

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(2) INFORMATION FOR SEQ ID NO:331:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:331:

TGATCTGTGG

(2) INFORMATION FOR SEQ ID NO:332:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:332:

GATCTGTGGA

(2) INFORMATION FOR SEQ ID NO:333:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:333:

ATCTGTGGAT

(2) INFORMATION FOR SEQ ID NO:334:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:334:

TCTGTGGATC

(2) INFORMATION FOR SEQ ID NO:335:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:335:

CTGTGGATCT

(2) INFORMATION FOR SEQ ID NO:336:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:336:

TGTGGATCTA

(2) INFORMATION FOR SEQ ID NO:337:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:337:

GTGGATCTAC

(2) INFORMATION FOR SEQ ID NO:338:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:338:

TGGATCTACC

(2) INFORMATION FOR SEQ ID NO:339:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:339:

GGATCTACCA

- 117 -

## (2) INFORMATION FOR SEQ ID NO:340:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:340:

GATCTACCAC

## (2) INFORMATION FOR SEQ ID NO:341:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:341:

ATCTACCACA

## (2) INFORMATION FOR SEQ ID NO:342:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:342:

TCTACCACAC

## (2) INFORMATION FOR SEQ ID NO:343:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:343:

CTACCACACA

## (2) INFORMATION FOR SEQ ID NO:344:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:344:

TACCACACAC

## (2) INFORMATION FOR SEQ ID NO:345:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:345:

ACCACACACA

## (2) INFORMATION FOR SEQ ID NO:346:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:346:

CCACACACAA

## (2) INFORMATION FOR SEQ ID NO:347:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:347:

CACACACAAG

## (2) INFORMATION FOR SEQ ID NO:348:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:348:

ACACACAAGG

## (2) INFORMATION FOR SEQ ID NO:349:

- 118 -

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:349:

CACACAAGGC

(2) INFORMATION FOR SEQ ID NO:350:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:350:

ACACAAGGCT

(2) INFORMATION FOR SEQ ID NO:351:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:351:

CACAAGGCTA

(2) INFORMATION FOR SEQ ID NO:352:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:352:

ACAAGGCTAC

(2) INFORMATION FOR SEQ ID NO:353:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:353:

CAAGGCTACT

(2) INFORMATION FOR SEQ ID NO:354:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:354:

AAGGCTACTT

(2) INFORMATION FOR SEQ ID NO:355:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:355:

AGGCTACTTC

(2) INFORMATION FOR SEQ ID NO:356:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:356:

GGCTACTTCC

(2) INFORMATION FOR SEQ ID NO:357:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:357:

GCTACTTCCC

(2) INFORMATION FOR SEQ ID NO:358:

- (i) SEQUENCE CHARACTERISTICS:

- 119 -

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:358:

CTACTTCCCT

(2) INFORMATION FOR SEQ ID NO:359:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:359:

TACTTCCCTG

(2) INFORMATION FOR SEQ ID NO:360:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:360:

ACTTCCCTGA

(2) INFORMATION FOR SEQ ID NO:361:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:361:

CTTCCCTGAT

(2) INFORMATION FOR SEQ ID NO:362:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:362:

TTCCTGATT

(2) INFORMATION FOR SEQ ID NO:363:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:363:

TCCCTGATTG

(2) INFORMATION FOR SEQ ID NO:364:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:364:

CCCTGATTGG

(2) INFORMATION FOR SEQ ID NO:365:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:365:

CCTGATTGGC

(2) INFORMATION FOR SEQ ID NO:366:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:366:

CTGATTGGCA

(2) INFORMATION FOR SEQ ID NO:367:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid

- 120 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:367:

TGATTGGCAG

(2) INFORMATION FOR SEQ ID NO:368:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:368:

GATTGGCAGA

(2) INFORMATION FOR SEQ ID NO:369:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:369:

ATTGGCAGAA

(2) INFORMATION FOR SEQ ID NO:370:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:370:

TTGGCAGAAC

(2) INFORMATION FOR SEQ ID NO:371:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:371:

TGGCAGAACT

(2) INFORMATION FOR SEQ ID NO:372:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:372:

GGCAGAACTA

(2) INFORMATION FOR SEQ ID NO:373:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:373:

GCAGAACTAC

(2) INFORMATION FOR SEQ ID NO:374:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:374:

CAGAACTACA

(2) INFORMATION FOR SEQ ID NO:375:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:375:

AGAACTACAC

(2) INFORMATION FOR SEQ ID NO:376:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 121 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:376:

GAACTACACA

(2) INFORMATION FOR SEQ ID NO:377:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:377:

AACTACACAC

(2) INFORMATION FOR SEQ ID NO:378:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:378:

ACTACACACC

(2) INFORMATION FOR SEQ ID NO:379:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:379:

CTACACACCA

(2) INFORMATION FOR SEQ ID NO:380:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:380:

TACACACCAG

(2) INFORMATION FOR SEQ ID NO:381:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:381:

ACACACCAGG

(2) INFORMATION FOR SEQ ID NO:382:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:382:

CACACCAGGG

(2) INFORMATION FOR SEQ ID NO:383:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:383:

ACACCAGGGC

(2) INFORMATION FOR SEQ ID NO:384:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:384:

CACCAGGGCC

(2) INFORMATION FOR SEQ ID NO:385:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 122 -

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:385:  
ACCAGGGCCA
- (2) INFORMATION FOR SEQ ID NO:386:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:386:  
CCAGGGCCAG
- (2) INFORMATION FOR SEQ ID NO:387:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:387:  
CAGGGCCAGG
- (2) INFORMATION FOR SEQ ID NO:388:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:388:  
AGGGCCAGGG
- (2) INFORMATION FOR SEQ ID NO:389:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:389:  
GGGCCAGGGG
- (2) INFORMATION FOR SEQ ID NO:390:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:390:  
GGCCAGGGGT
- (2) INFORMATION FOR SEQ ID NO:391:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:391:  
GCCAGGGGTC
- (2) INFORMATION FOR SEQ ID NO:392:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:392:  
CCAGGGGTCA
- (2) INFORMATION FOR SEQ ID NO:393:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:393:  
CAGGGGTCAG
- (2) INFORMATION FOR SEQ ID NO:394:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:394:



- 123 -

AGGGGTCAGA

## (2) INFORMATION FOR SEQ ID NO:395:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:395:

GGGGTCAGAT

## (2) INFORMATION FOR SEQ ID NO:396:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:396:

GGGTCAGATA

## (2) INFORMATION FOR SEQ ID NO:397:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:397:

GGTCAGATAT

## (2) INFORMATION FOR SEQ ID NO:398:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:398:

GTCAGATATC

## (2) INFORMATION FOR SEQ ID NO:399:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:399:

TCAGATATCC

## (2) INFORMATION FOR SEQ ID NO:400:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:400:

CAGATATCCA

## (2) INFORMATION FOR SEQ ID NO:401:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:401:

AGATATCCAC

## (2) INFORMATION FOR SEQ ID NO:402:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:402:

GATATCCACT

## (2) INFORMATION FOR SEQ ID NO:403:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:403:

ATATCCACTG

- 124 -

## (2) INFORMATION FOR SEQ ID NO:404:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:404:

TATCCACTGA

## (2) INFORMATION FOR SEQ ID NO:405:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:405:

ATCCACTGAC

## (2) INFORMATION FOR SEQ ID NO:406:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:406:

TCCACTGACC

## (2) INFORMATION FOR SEQ ID NO:407:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:407:

CCTGACCT

## (2) INFORMATION FOR SEQ ID NO:408:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:408:

CACTGACCTT

## (2) INFORMATION FOR SEQ ID NO:409:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:409:

ACTGACCTTT

## (2) INFORMATION FOR SEQ ID NO:410:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:410:

CTGACCTTTG

## (2) INFORMATION FOR SEQ ID NO:411:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:411:

TGACCTTTGG

## (2) INFORMATION FOR SEQ ID NO:412:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:412:

GACCTTTGGA

## (2) INFORMATION FOR SEQ ID NO:413:

- 125 -

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:413:

ACCTTTGGAT

(2) INFORMATION FOR SEQ ID NO:414:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:414:

CCTTTGGATG

(2) INFORMATION FOR SEQ ID NO:415:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:415:

CTTTGGATGG

(2) INFORMATION FOR SEQ ID NO:416:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:416:

TTTGGATGGT

(2) INFORMATION FOR SEQ ID NO:417:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:417:

TTGGATGGTG

(2) INFORMATION FOR SEQ ID NO:418:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:418:

TGGATGGTGC

(2) INFORMATION FOR SEQ ID NO:419:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:419:

GGATGGTGCT

(2) INFORMATION FOR SEQ ID NO:420:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:420:

GATGGTGCTA

(2) INFORMATION FOR SEQ ID NO:421:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:421:

ATGGTGCTAC

(2) INFORMATION FOR SEQ ID NO:422:

- (i) SEQUENCE CHARACTERISTICS:

- 126 -

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:422:

TGGTGCTACA

(2) INFORMATION FOR SEQ ID NO:423:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:423:

GGTGCTACAA

(2) INFORMATION FOR SEQ ID NO:424:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:424:

GTGCTACAAG

(2) INFORMATION FOR SEQ ID NO:425:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:425:

TGCTACAAGC

(2) INFORMATION FOR SEQ ID NO:426:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:426:

GCTACAAGCT

(2) INFORMATION FOR SEQ ID NO:427:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:427:

CTACAAGCTA

(2) INFORMATION FOR SEQ ID NO:428:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:428:

TACAAGCTAG

(2) INFORMATION FOR SEQ ID NO:429:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:429:

ACAAGCTAGT

(2) INFORMATION FOR SEQ ID NO:430:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:430:

CAAGCTAGTA

(2) INFORMATION FOR SEQ ID NO:431:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid

- 127 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:431:

AAGCTAGTAC

(2) INFORMATION FOR SEQ ID NO:432:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:432:

AGCTAGTACC

(2) INFORMATION FOR SEQ ID NO:433:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:433:

GCTAGTACCA

(2) INFORMATION FOR SEQ ID NO:434:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:434:

CTAGTACCAG

(2) INFORMATION FOR SEQ ID NO:435:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:435:

TAGTACCAGT

(2) INFORMATION FOR SEQ ID NO:436:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:436:

AGTACCAGTT

(2) INFORMATION FOR SEQ ID NO:437:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:437:

GTACCAGTTG

(2) INFORMATION FOR SEQ ID NO:438:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:438:

TACCAGTTGA

(2) INFORMATION FOR SEQ ID NO:439:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:439:

ACCAGTTGAG

(2) INFORMATION FOR SEQ ID NO:440:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 128 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:440:

CCAGTTGAGC

(2) INFORMATION FOR SEQ ID NO:441:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:441:

CAGTTGAGCC

(2) INFORMATION FOR SEQ ID NO:442:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:442:

AGTTGAGCCA

(2) INFORMATION FOR SEQ ID NO:443:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:443:

GTTGAGCCAG

(2) INFORMATION FOR SEQ ID NO:444:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:444:

TTGAGCCAGA

(2) INFORMATION FOR SEQ ID NO:445:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:445:

TGAGCCAGAT

(2) INFORMATION FOR SEQ ID NO:446:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:446:

GAGCCAGATA

(2) INFORMATION FOR SEQ ID NO:447:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:447:

AGCCAGATAA

(2) INFORMATION FOR SEQ ID NO:448:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:448:

GCCAGATAAG

(2) INFORMATION FOR SEQ ID NO:449:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 129 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:449:

CCAGATAAGG

(2) INFORMATION FOR SEQ ID NO:450:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:450:

CAGATAAGGT

(2) INFORMATION FOR SEQ ID NO:451:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:451:

AGATAAGGTA

(2) INFORMATION FOR SEQ ID NO:452:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:452:

GATAAGGTAG

(2) INFORMATION FOR SEQ ID NO:453:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:453:

ATAAGGTAGA

(2) INFORMATION FOR SEQ ID NO:454:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:454:

TAAGGTAGAA

(2) INFORMATION FOR SEQ ID NO:455:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:455:

AAGGTAGAAG

(2) INFORMATION FOR SEQ ID NO:456:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:456:

AGGTAGAAGA

(2) INFORMATION FOR SEQ ID NO:457:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:457:

GGTAGAAGAG

(2) INFORMATION FOR SEQ ID NO:458:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:458:

- 130 -

GTAGAAGAGG

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(2) INFORMATION FOR SEQ ID NO:459:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:459:

TAGAAGAGGC

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:463:

AGAGGCCAAT

(2) INFORMATION FOR SEQ ID NO:464:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:464:

GAGGCCAATA

(2) INFORMATION FOR SEQ ID NO:460:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:460:

AGAAGAGGCC

(2) INFORMATION FOR SEQ ID NO:465:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:465:

AGGCCAATAA

(2) INFORMATION FOR SEQ ID NO:461:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:461:

GAAGAGGCCA

(2) INFORMATION FOR SEQ ID NO:466:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:466:

GGCCAATAAA

(2) INFORMATION FOR SEQ ID NO:462:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:462:

AAGAGGCCAA

(2) INFORMATION FOR SEQ ID NO:467:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:467:

GCCAATAAAG

(2) INFORMATION FOR SEQ ID NO:463:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid



- 131 -

## (2) INFORMATION FOR SEQ ID NO:468:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:468:

CCAATAAAGG

## (2) INFORMATION FOR SEQ ID NO:469:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:469:

CAATAAAGGA

## (2) INFORMATION FOR SEQ ID NO:470:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:470:

AATAAAGGAG

## (2) INFORMATION FOR SEQ ID NO:471:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:471:

ATAAAGGAGA

## (2) INFORMATION FOR SEQ ID NO:472:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:472:

TAAAGGAGAG

## (2) INFORMATION FOR SEQ ID NO:473:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:473:

AAAGGAGAGA

## (2) INFORMATION FOR SEQ ID NO:474:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:474:

AAGGAGAGAA

## (2) INFORMATION FOR SEQ ID NO:475:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:475:

AGGAGAGAAC

## (2) INFORMATION FOR SEQ ID NO:476:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:476:

GGAGAGAACA

## (2) INFORMATION FOR SEQ ID NO:477:

- 132 -

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:477:

GAGAGAACAC

(2) INFORMATION FOR SEQ ID NO:478:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:478:

AGAGAACACC

(2) INFORMATION FOR SEQ ID NO:479:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:479:

GAGAACACCA

(2) INFORMATION FOR SEQ ID NO:480:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:480:

AGAACACCAG

(2) INFORMATION FOR SEQ ID NO:481:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:481:

GAACACCAGC

(2) INFORMATION FOR SEQ ID NO:482:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:482:

AACACCAGCT

(2) INFORMATION FOR SEQ ID NO:483:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:483:

ACACCAGCTT

(2) INFORMATION FOR SEQ ID NO:484:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:484:

CACCAGCTTG

(2) INFORMATION FOR SEQ ID NO:485:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:485:

ACCAGCTTGT

(2) INFORMATION FOR SEQ ID NO:486:

- (i) SEQUENCE CHARACTERISTICS:

- 133 -

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:486:

CCAGCTTGTT

(2) INFORMATION FOR SEQ ID NO:487:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:487:

CAGCTTGTTA

(2) INFORMATION FOR SEQ ID NO:488:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:488:

AGCTTGTTAC

(2) INFORMATION FOR SEQ ID NO:489:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:489:

GCTTGTTACA

(2) INFORMATION FOR SEQ ID NO:490:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:490:

CTTGTTACAC

(2) INFORMATION FOR SEQ ID NO:491:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:491:

TTGTTACACC

(2) INFORMATION FOR SEQ ID NO:492:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:492:

TGTTACACCC

(2) INFORMATION FOR SEQ ID NO:493:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:493:

GTTACACCCCT

(2) INFORMATION FOR SEQ ID NO:494:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:494:

TTACACCCCTG

(2) INFORMATION FOR SEQ ID NO:495:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid

- 134 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:495:

TACACCCTGT

(2) INFORMATION FOR SEQ ID NO:496:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:496:

ACACCCTGTG

(2) INFORMATION FOR SEQ ID NO:497:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:497:

CACCCTGTGA

(2) INFORMATION FOR SEQ ID NO:498:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:498:

ACCCTGTGAG

(2) INFORMATION FOR SEQ ID NO:499:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:499:

CCCTGTGAGC

(2) INFORMATION FOR SEQ ID NO:500:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:500:

CCTGTGAGCC

(2) INFORMATION FOR SEQ ID NO:501:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:501:

CTGTGAGCCT

(2) INFORMATION FOR SEQ ID NO:502:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:502:

TGTGAGCCTG

(2) INFORMATION FOR SEQ ID NO:503:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:503:

GTGAGCCTGC

(2) INFORMATION FOR SEQ ID NO:504:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 135 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:504:

TGAGCCTGCA

(2) INFORMATION FOR SEQ ID NO:505:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:505:

GAGCCTGCAT

(2) INFORMATION FOR SEQ ID NO:506:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:506:

AGCCTGCATG

(2) INFORMATION FOR SEQ ID NO:507:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:507:

GCCTGCATGG

(2) INFORMATION FOR SEQ ID NO:508:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:508:

CCTGCATGGA

(2) INFORMATION FOR SEQ ID NO:509:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:509:

CTGCATGGAA

(2) INFORMATION FOR SEQ ID NO:510:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:510:

TGCATGGAAT

(2) INFORMATION FOR SEQ ID NO:511:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:511:

GCATGGAATG

(2) INFORMATION FOR SEQ ID NO:512:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:512:

CATGGAATGG

(2) INFORMATION FOR SEQ ID NO:513:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 136 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:513:  
ATGGAATGGA

(2) INFORMATION FOR SEQ ID NO:514:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:514:  
TGGAATGGAT

(2) INFORMATION FOR SEQ ID NO:515:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:515:  
GGAATGGATG

(2) INFORMATION FOR SEQ ID NO:516:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:516:  
GAATGGATGA

(2) INFORMATION FOR SEQ ID NO:517:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:517:  
AATGGATGAC

(2) INFORMATION FOR SEQ ID NO:518:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:518:  
ATGGATGACC

(2) INFORMATION FOR SEQ ID NO:519:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:519:  
TGGATGACCC

(2) INFORMATION FOR SEQ ID NO:520:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:520:  
GGATGACCCCT

(2) INFORMATION FOR SEQ ID NO:521:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:521:  
GATGACCCCTG

(2) INFORMATION FOR SEQ ID NO:522:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:522:

- 137 -

ATGACCOCTGA

## (2) INFORMATION FOR SEQ ID NO:523:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:523:

TGACCCTGAG

## (2) INFORMATION FOR SEQ ID NO:524:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:524:

GACCCTGAGA

## (2) INFORMATION FOR SEQ ID NO:525:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:525:

ACCCTGAGAG

## (2) INFORMATION FOR SEQ ID NO:526:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:526:

CCCTGAGAGA

## (2) INFORMATION FOR SEQ ID NO:527:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:527:

CCTGAGAGAG

## (2) INFORMATION FOR SEQ ID NO:528:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:528:

CTGAGAGAGA

## (2) INFORMATION FOR SEQ ID NO:529:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:529:

TGAGAGAGAA

## (2) INFORMATION FOR SEQ ID NO:530:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:530:

GAGAGAGAAG

## (2) INFORMATION FOR SEQ ID NO:531:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:531:

AGAGAGAAGT

- 138 -

## (2) INFORMATION FOR SEQ ID NO:532:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:532:

GAGAGAAGTG

## (2) INFORMATION FOR SEQ ID NO:533:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:533:

AGAGAAAGTGT

## (2) INFORMATION FOR SEQ ID NO:534:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:534:

GAGAAGTGTT

## (2) INFORMATION FOR SEQ ID NO:535:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:535:

AGAAGTGTTA

## (2) INFORMATION FOR SEQ ID NO:536:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:536:

GAAGTGTTAG

## (2) INFORMATION FOR SEQ ID NO:537:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:537:

AAGTGTTAGA

## (2) INFORMATION FOR SEQ ID NO:538:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:538:

AGTGTTAGAG

## (2) INFORMATION FOR SEQ ID NO:539:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:539:

GTGTTAGAGT

## (2) INFORMATION FOR SEQ ID NO:540:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:540:

TGTTAGAGTG

## (2) INFORMATION FOR SEQ ID NO:541:



- 139 -

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:541:

GTTAGAGTGG

(2) INFORMATION FOR SEQ ID NO:542:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:542:

TTAGAGTGGG

(2) INFORMATION FOR SEQ ID NO:543:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:543:

TAGAGTGGAG

(2) INFORMATION FOR SEQ ID NO:544:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:544:

AGAGTGGAGG

(2) INFORMATION FOR SEQ ID NO:545:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:545:

GAGTGGAGGT

(2) INFORMATION FOR SEQ ID NO:546:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:546:

AGTGGAGGTT

(2) INFORMATION FOR SEQ ID NO:547:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:547:

GTGGAGGTTT

(2) INFORMATION FOR SEQ ID NO:548:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:548:

TGGAGGTTTG

(2) INFORMATION FOR SEQ ID NO:549:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:549:

GGAGGTTTGA

(2) INFORMATION FOR SEQ ID NO:550:

- (i) SEQUENCE CHARACTERISTICS:

- 140 -

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:550:

GAGGTTTGAC

(2) INFORMATION FOR SEQ ID NO:551:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:551:

AGGTTTGACA

(2) INFORMATION FOR SEQ ID NO:552:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:552:

GGTTTGACAG

(2) INFORMATION FOR SEQ ID NO:553:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:553:

GTTTGACAGC

(2) INFORMATION FOR SEQ ID NO:554:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:554:

TTTGACAGCC

(2) INFORMATION FOR SEQ ID NO:555:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:555:

TTGACAGCCG

(2) INFORMATION FOR SEQ ID NO:556:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:556:

TGACAGCCGC

(2) INFORMATION FOR SEQ ID NO:557:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:557:

GACAGCCGCC

(2) INFORMATION FOR SEQ ID NO:558:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:558:

ACAGCCGCCT

(2) INFORMATION FOR SEQ ID NO:559:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid

- 141 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:559:

CAGCCGCCTA

(2) INFORMATION FOR SEQ ID NO:560:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:560:

AGCCGCCTAG

(2) INFORMATION FOR SEQ ID NO:561:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:561:

GCCGCCTAGC

(2) INFORMATION FOR SEQ ID NO:562:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:562:

CCGCCTAGCA

(2) INFORMATION FOR SEQ ID NO:563:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:563:

CGCCTAGCAT

(2) INFORMATION FOR SEQ ID NO:564:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:564:

GCCTAGCATT

(2) INFORMATION FOR SEQ ID NO:565:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:565:

CCTAGCATTT

(2) INFORMATION FOR SEQ ID NO:566:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:566:

CTAGCATTTT

(2) INFORMATION FOR SEQ ID NO:567:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:567:

TAGCATTTCA

(2) INFORMATION FOR SEQ ID NO:568:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 142 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:568:

AGCATTTCAT

(2) INFORMATION FOR SEQ ID NO:569:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:569:

GCATTTCATC

(2) INFORMATION FOR SEQ ID NO:570:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:570:

CATTTTCATCA

(2) INFORMATION FOR SEQ ID NO:571:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:571:

ATTTCATCAC

(2) INFORMATION FOR SEQ ID NO:572:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:572:

TTTCATCACG

(2) INFORMATION FOR SEQ ID NO:573:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:573:

TTCATCACGT

(2) INFORMATION FOR SEQ ID NO:574:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:574:

TCATCACGTG

(2) INFORMATION FOR SEQ ID NO:575:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:575:

CATCACGTGG

(2) INFORMATION FOR SEQ ID NO:576:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:576:

ATCACGTGGC

(2) INFORMATION FOR SEQ ID NO:577:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 143 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:577:

TCACGTGGCC

(2) INFORMATION FOR SEQ ID NO:578:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:578:

CACGTGGCCC

(2) INFORMATION FOR SEQ ID NO:579:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:579:

ACGTGGCCCG

(2) INFORMATION FOR SEQ ID NO:580:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:580:

CGTGGCCCGA

(2) INFORMATION FOR SEQ ID NO:581:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:581:

GTGGCCCGAG

(2) INFORMATION FOR SEQ ID NO:582:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:582:

TGGCCCGAGA

(2) INFORMATION FOR SEQ ID NO:583:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:583:

GGCCCGAGAG

(2) INFORMATION FOR SEQ ID NO:584:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:584:

GCCCGAGAGC

(2) INFORMATION FOR SEQ ID NO:585:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:585:

CCCGAGAGCT

(2) INFORMATION FOR SEQ ID NO:586:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:586:

- 144 -

CCGAGAGCTG

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(2) INFORMATION FOR SEQ ID NO:587:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:587:

CGAGAGCTGC

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:591:

AGCTGCATCC

(2) INFORMATION FOR SEQ ID NO:592:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:592:

GCTGCATCCG

(2) INFORMATION FOR SEQ ID NO:588:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:588:

GAGAGCTGCA

(2) INFORMATION FOR SEQ ID NO:593:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:593:

CTGCATCCGG

(2) INFORMATION FOR SEQ ID NO:589:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:589:

AGAGCTGCAT

(2) INFORMATION FOR SEQ ID NO:594:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:594:

TGATCCGGA

(2) INFORMATION FOR SEQ ID NO:590:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:590:

GAGCTGCATC

(2) INFORMATION FOR SEQ ID NO:595:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:595:

GCATCCGGAG

(2) INFORMATION FOR SEQ ID NO:591:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid

- 145 -

## (2) INFORMATION FOR SEQ ID NO:596:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:596:

CATCCGGAGT

## (2) INFORMATION FOR SEQ ID NO:597:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:597:

ATCCGGAGTA

## (2) INFORMATION FOR SEQ ID NO:598:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:598:

TCCGGAGTAC

## (2) INFORMATION FOR SEQ ID NO:599:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:599:

CCGGAGTACT

## (2) INFORMATION FOR SEQ ID NO:600:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:600:

CGGAGTACTT

## (2) INFORMATION FOR SEQ ID NO:601:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:601:

GGAGTACTTC

## (2) INFORMATION FOR SEQ ID NO:602:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:602:

GAGTACTTCA

## (2) INFORMATION FOR SEQ ID NO:603:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:603:

AGTACTTCAA

## (2) INFORMATION FOR SEQ ID NO:604:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:604:

GTACTTCAAG

## (2) INFORMATION FOR SEQ ID NO:605:

- 146 -

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:605:

TACTTCAAGA

(2) INFORMATION FOR SEQ ID NO:606:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:606:

ACTTCAAGAA

(2) INFORMATION FOR SEQ ID NO:607:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:607:

CTTCAAGAAC

(2) INFORMATION FOR SEQ ID NO:608:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:608:

TTCAAGAACT

(2) INFORMATION FOR SEQ ID NO:609:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:609:

TCAAGAACTG

(2) INFORMATION FOR SEQ ID NO:610:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:610:

CAAGAACTGC

(2) INFORMATION FOR SEQ ID NO:611:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:611:

AAGAACTGCT

(2) INFORMATION FOR SEQ ID NO:612:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:612:

AGAACTGCTG

(2) INFORMATION FOR SEQ ID NO:613:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single



- 147 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:613:

GAACTGCTGA

- 148 -

## (2) INFORMATION FOR SEQ ID NO:614:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1305 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:614:

GAAGAGATTT GGGAGAACAT GACCTGGATG CAGTGGGAAA AAGAAATTCA CAATCACACA	60
AAATACATAT ACTCCTTACT TGAAAAATCG CAGAACCAAC AAGAAAAGAA TGAACAAGAA	120
CTATTGGAAT TGGATCAATG GGCAAGTTTG TGAATTGGT TTGACATAAC AAAATGGCTG	180
TGGTATATAA AAATATTCAT AATGGTAGTA GGAGGCTTGA TAGGTTTAAG AATAGTTTTT	240
GCTGTACTTT CTATAGTGAA TAGAGTTAGG CAGGGATACT CACCATTGTC GTTTCAGACC	300
CTCCTCCCAA CCCCAGGGG ACCCGACAGG CCCGAAGGAA TCGAAGAAGA AGGTGGAGAG	360
AGAGACAGAG ACAGATCCAC TCGATTAGTA CACGGATTCT TAGCACTTTT CTGGGACGAC	420
CTGAGGAGCC TGTGCCTCTT CCTCTACCAC CACTTGAGAG ACTTACTCTT GATTGTAACA	480
AGGATTGTGG AACTTCTGGG ACGCAGGGGA TGGGAAGCCC TCAAATATTG GTGGAACCTC	540
CTAAAGTATT GGAGCCAGGA ACTGCAGAAG AGTGTCTGTTA TCTTGCTCAA TGCCACCGCC	600
ATAGCAGTAG CTGAGGGGAC AGATAGAGTT TTAGAAGTAT TACAAAGAGC TTATAGAGCT	660
ATCCTCCACA TACCTAGAAG AATAAGACAG GGCCTCGAAA TGGCTTTGCT ATAAAATGGG	720
TGGCAAGTGA GCAAAAAGTA GTGTAGTCAG ATAGCATGCA TCATAAGGGG TGGGGGCCAA	780
CAACTAACAA TGCTGATCGT GCCTGGCTAG AAGCACAAGA GAAGGAAGAA GCGGGTTTTC	840
CAGTCAAACC TCAGGTAGCT GTAGATCTTA GCCACTTTTT AAAAGAAAAG GGGGGACTGG	900
AAGGGCTAAT TCACTCCCAA AGAAGACAAG ATACACAGTG CTGCAAACTA TTACCAGTGG	960
AGTCAGCGAA GATAGAAGAG GCCAATGGAG GAGAAAACCA CAGATTGTTT TGTGGGGAC	1020
TTTCCATCCG TTGGGGACTT TCCAAGGCGG CGTGGCCTGG GTGACTAGTT CCGGTGGGGA	1080
CTTTCCAAGA AGGCGCGGCC TGGGCGGGAC TGGGGAGTGG CGAGCCCTCA GATGCTGCAT	1140
ATAAGCAGCT GCTTTCTGCT GTTACTGGGT CTCTCGGGTT AGACCAGATC TGAGCCTGGG	1200
AGCTCTCTGG CTAAGTAGG AACCCTGCT TTAAGCCTCA ATAAAGCTTG CCTTGAGTGC	1260
TTCAAGTAGT GTGTGCCCGT CTGTTGTGTG ACTCTGGTAT CTAGA	1305

- 149 -

## (2) INFORMATION FOR SEQ ID NO:615:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1208 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:615:

GAAACAATTT GGGATAACAT GACCTGGATG CAGTGGGAAA GAGAAATTGA CAATTACACA	60
AACATAATAT ACACCTTAAT TGAAGAATCG CAGAACCAAC AAGAAAAAAA TGAAC TAGAA	120
TTATTGGAAT TGGATAAATG GGCAAATTTG TGAATTGGT TTAGTATATC AACTGGCTA	180
TGGTATATAA AATTATTCAT AATGGTAGTA GGAGGCTTGG TAGGTTTAAG AATAGTTTTT	240
ACTGTACTTT CTATAGTGAA TAGAGTTAGG CAGGGATACT CACCATTGTC GTTTCAGACC	300
CACCTCCCAA CCCCGAAGGG ACCCGACAGG CCAGAAGGAA TCGAAGAAGA AGGTGGAGAG	360
AGAGACAGAG GCAGCTCCAC TCGATTAGTG CACGGATTCT TAGCACITTT CTGGGACGAC	420
CTGAGGAGTC TGTGCCTCTT CAGCTACCAC CACTTGAGAG ACTTACTCTT GATTGTAACG	480
AGGATTGTGG AACTTCTGGG ACGCAGGGGA TGGGAAGCCC TCAAATACTG GTGGAATCTC	540
CTGCAGTATT GGAGGCAGGA ACTACAGAAG AGTGCTGTTA GCTTGTTCAA TGGCACGGCC	600
ATAGCAGTAG CTGAGGGGAC AGATAGAGTT ATAGAAGCTT TACGAAGGGC TTATAGAGCT	660
ATTCTCCACA TACCTAGAAG AATAAGACAG GGCTTAGAAA GGGCTTTGCT ATAAAATGGG	720
TGGCAAGTGG TCAGAAAGTA GTGTGGTTAG AAGGCATGTA CCTTTAAGAC AAGGCAGCTA	780
TAGATCTTAG CCGCTTTTAA AAAGAAAAGG GGGGACTGGA AGGGCTAATT CACTCACAGA	840
GAAGATCAGT TGAACCAGAA GAAGATAGAA GAGGCCATGA AGAAGAAAAC AACAGATTGT	900
TCCGTTTGTT CCGTTGGGGA CTTTCAGGA GACGTGGCCT GAGTGATAAG CCGCTGGGGA	960
CTTCCGAAG AGGCGTGACG GGACTTTCCA AGGCGACGTG GCCTGGGCGG GACTGGGGAG	1020
TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTCT GCCTGTACTG GGTCTCTCTG	1080
GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA GGAACCCAC TGCTTAAGCC	1140
TCAATAAAGC TTGCCTTGAG TGCTTCAAGT AGTGTGTGCC CGTCTGTTGT GTGACTCTGG	1200
TATCTAGA	1208

- 150 -

## (2) INFORMATION FOR SEQ ID NO:616:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:616:

TGGAAGGGCT AATTGGT

18

## (2) INFORMATION FOR SEQ ID NO:617:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:617:

ATCTTCCTA AAAAATTAGC CTGTC

25

## (2) INFORMATION FOR SEQ ID NO:618:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:618:

AGGCTCAGAT CTGGTCTAAC

20

## (2) INFORMATION FOR SEQ ID NO:619:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:619:

AGCAGCAGGA AGCACTATGG

20

- 151 -

## (2) INFORMATION FOR SEQ ID NO:620:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:620:

TGCTAGAGAT TTTCCACAC

19

## (2) INFORMATION FOR SEQ ID NO:621:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:621:

AGTGAATAGA GTTAGGCAGG

20

## (2) INFORMATION FOR SEQ ID NO:622:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:622:

GTAAGACAGT ATGATCAGAT A

21

## (2) INFORMATION FOR SEQ ID NO:623:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:623:

TTGTAGGGAA TTCCAAATTC C

21

- 152 -

## (2) INFORMATION FOR SEQ ID NO:624:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:624:

CAGGATCCTA CACCTGTCAA CATAAT

26

## (2) INFORMATION FOR SEQ ID NO:625:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:625:

GGGAATTCTT TATTCCTGCT TG

22

## (2) INFORMATION FOR SEQ ID NO:626:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:626:

CCAGAAGTTC CACAATCC

18

## (2) INFORMATION FOR SEQ ID NO:627:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:627:

TTCTTCTAGG TATGTGGAG

19

- 153 -

## (2) INFORMATION FOR SEQ ID NO:628:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:628:

AGTGAATTAG CCCTTCCAG

19

## (2) INFORMATION FOR SEQ ID NO:629:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:629:

TGCTAGAGAT TTTCACAC

19

## (2) INFORMATION FOR SEQ ID NO:630:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:630:

TGCTCTGGAA AACTCAT

17

## (2) INFORMATION FOR SEQ ID NO:631:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:631:

CTTCTATAG TGAATAGAG

19

- (11) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:632;

17

(2) INFORMATION FOR SEQ ID NO:633:

- (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:633:

17

(2) INFORMATION FOR SEQ ID NO:634:

- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:634:

INFORMATION FOR SEQ ID NO:635:

- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:635:

Asp Pro His  
35



- 155 -

## (2) INFORMATION FOR SEQ ID NO:636:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 36 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:636:

```

Pro Ser Ser Gln Pro Arg Gly Asp Pro Thr Gly Pro Lys Glu Ser Lys
1           5           10           15

Lys Lys Val Glu Arg Glu Thr Glu Thr Asp Pro Leu Asp Tyr Thr
                20           25           30

Asp Ser His
                35

```

## (2) INFORMATION FOR SEQ ID NO:637:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 36 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:637:

```

Pro Thr Ser Gln Pro Arg Arg Asp Pro Thr Gly Gln Lys Glu Ser Lys
1           5           10           15

Lys Lys Val Glu Arg Glu Thr Glu Ala Ala Pro Leu Asp Cys Thr
                20           25           30

Asp Ser His
                35

```

## (2) INFORMATION FOR SEQ ID NO:638:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 110 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:638:

```

Asp Pro Pro Pro Asn Pro Glu Gly Thr Arg Gln Ala Arg Arg Asn Arg
1           5           10           15

Arg Arg Arg Trp Arg Glu Arg Gln Arg Gln Ile His Ser Ile Ser Glu
                20           25           30

Arg Ile Leu Ser Thr Tyr Leu Gly Arg Ser Ala Glu Pro Val Pro Leu
                35           40           45

Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Asp Cys Asn Glu Asp Cys
                50           55           60

Gly Thr Ser Gly Thr Gln Gly Val Gly Ser Pro Gln Ile Leu Val Glu
65           70           75           80

```

- 156 -

Ser Pro Thr Val Leu Glu Ser Gly Thr Lys Glu Cys Cys Leu 95  
                                     85                                    90  
 Ala Gln Cys His Ser His Ser Ser Ser Gly Asp Arg 110  
                                     100                                    105                                    110

## (2) INFORMATION FOR SEQ ID NO:639:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 110 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:639:

Asp Pro Pro Pro Asn Pro Glu Gly Thr Arg Gln Ala Arg Arg Asn Arg 15  
   1                                    5                                    10                                    15  
 Arg Arg Arg Trp Arg Glu Arg Gln Arg Gln Ile His Ser Ile Ser Thr 30  
                                     20                                    25                                    30  
 Arg Ile Leu Ser Thr Phe Leu Gly Arg Pro Glu Glu Pro Val Pro Leu 45  
                                     35                                    40                                    45  
 Pro Leu Pro Pro Leu Glu Arg Leu Thr Leu Asp Cys Asn Lys Asp Cys 60  
                                     50                                    55                                    60  
 Gly Thr Ser Gly Thr Gln Gly Met Gly Ser Pro Gln Ile Leu Val Glu 80  
   65                                    70                                    75                                    80  
 Pro Pro Lys Val Leu Glu Pro Gly Thr Ala Glu Glu Cys Cys Tyr Leu 95  
                                     85                                    90                                    95  
 Ala Gln Cys His Arg His Ser Ser Ser Gly Asp Arg 110  
                                     100                                    105                                    110

## (2) INFORMATION FOR SEQ ID NO:640:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 110 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:640:

Asp Pro Pro Pro Asn Pro Glu Gly Thr Arg Gln Ala Arg Arg Asn Arg 15  
   1                                    5                                    10                                    15  
 Arg Arg Arg Trp Arg Glu Arg Gln Arg Gln Leu His Ser Ile Ser Ala 30  
                                     20                                    25                                    30  
 Arg Ile Leu Ser Thr Phe Leu Gly Arg Pro Glu Glu Ser Val Pro Leu 45  
                                     35                                    40                                    45  
 Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Asp Cys Asn Glu Asp Cys 60  
                                     50                                    55                                    60  
 Gly Thr Ser Gly Thr Gln Gly Met Gly Ser Pro Gln Ile Leu Val Glu 80  
   65                                    70                                    75                                    80  
 Ser Pro Ala Val Leu Glu Ala Gly Thr Thr Glu Glu Cys Cys Leu 95  
                                     85                                    90                                    95

- 157 -

Val Gln Trp His Gly His Ser Ser Ser Gly Asp Arg  
 100 105 110

## (2) INFORMATION FOR SEQ ID NO:641:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 237 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:641:

Glu Gln Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile  
 1 5 10 15  
 Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn  
 20 25 30  
 Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala  
 35 40 45  
 Ser Leu Trp Asn Trp Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys  
 50 55 60  
 Leu Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe  
 65 70 75 80  
 Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu  
 85 90 95  
 Ser Phe Gln Thr His Leu Pro Ile Pro Arg Gly Pro Asp Arg Pro Glu  
 100 105 110  
 Gly Ile Glu Glu Glu Gly Gly Glu Arg Asp Arg Asp Arg Ser Ile Arg  
 115 120 125  
 Leu Val Asn Gly Ser Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu  
 130 135 140  
 Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr  
 145 150 155 160  
 Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr  
 165 170 175  
 Trp Trp Asn Leu Leu Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala  
 180 185 190  
 Val Asn Leu Leu Asn Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp  
 195 200 205  
 Arg Val Ile Glu Val Leu Gln Ala Ala Tyr Arg Ala Ile Arg His Ile  
 210 215 220  
 Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ile Leu Leu  
 225 230 235

- 158 -

## (2) INFORMATION FOR SEQ ID NO:642:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 237 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:642:

Glu Glu Ile Trp Glu Asn Met Thr Trp Met Gln Trp Glu Lys Glu Ile  
 1 5 10 15  
 His Asn His Thr Lys Tyr Ile Tyr Ser Leu Leu Glu Lys Ser Gln Asn  
 20 25 30  
 Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Gln Trp Ala  
 35 40 45  
 Ser Leu Trp Asn Trp Phe Asp Ile Thr Lys Trp Leu Trp Tyr Ile Lys  
 50 55 60  
 Ile Phe Ile Met Val Val Gly Gly Leu Ile Gly Leu Arg Ile Val Phe  
 65 70 75 80  
 Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu  
 85 90 95  
 Ser Phe Gln Thr Leu Leu Pro Thr Pro Arg Gly Pro Asp Arg Pro Glu  
 100 105 110  
 Gly Ile Glu Glu Met Gly Gly Glu Arg Asp Arg Asp Arg Ser Thr Arg  
 115 120 125  
 Leu Val His Gly Phe Leu Ala Leu Phe Trp Asp Asp Leu Arg Ser Leu  
 130 135 140  
 Cys Leu Phe Leu Tyr His His Leu Arg Asp Leu Leu Leu Ile Val Thr  
 145 150 155 160  
 Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr  
 165 170 175  
 Trp Trp Asn Leu Leu Lys Tyr Trp Ser Gln Glu Leu Gln Lys Ser Ala  
 180 185 190  
 Val Ile Leu Leu Asn Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp  
 195 200 205  
 Arg Val Leu Glu Val Leu Gln Arg Ala Tyr Arg Ala Ile Leu His Ile  
 210 215 220  
 Pro Arg Arg Ile Arg Gln Gly Leu Glu Met Ala Leu Leu  
 225 230 235

- 159 -

## (2) INFORMATION FOR SEQ ID NO:643:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 237 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:643:

Glu Thr Ile Trp Asp Asn Met Thr Trp Met Gln Trp Glu Arg Glu Ile  
 1 5 10 15  
 Asp Asn Tyr Thr Asn Ile Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn  
 20 25 30  
 Gln Gln Glu Lys Asn Glu Leu Glu Leu Leu Glu Leu Asp Lys Trp Ala  
 35 40 45  
 Asn Leu Trp Asn Trp Phe Ser Ile Ser Asn Trp Leu Trp Tyr Ile Lys  
 50 55 60  
 Leu Phe Ile Met Val Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe  
 65 70 75 80  
 Thr Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu  
 85 90 95  
 Ser Phe Gln Thr His Leu Pro Thr Pro Lys Gly Pro Asp Arg Pro Glu  
 100 105 110  
 Gly Ile Glu Glu Glu Gly Gly Glu Arg Asp Arg Gly Ser Ser Thr Arg  
 115 120 125  
 Leu Val His Gly Phe Leu Ala Leu Phe Trp Asp Asp Leu Arg Ser Leu  
 130 135 140  
 Cys Leu Phe Ser Tyr His His Leu Arg Asp Leu Leu Leu Ile Val Thr  
 145 150 155 160  
 Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr  
 165 170 175  
 Trp Trp Asn Leu Leu Gln Tyr Trp Arg Gln Glu Leu Gln Lys Ser Ala  
 180 185 190  
 Val Ser Leu Phe Asn Gly Thr Ala Ile Ala Val Ala Glu Gly Thr Asp  
 195 200 205  
 Arg Val Ile Glu Ala Leu Arg Arg Ala Tyr Arg Ala Ile Leu His Ile  
 210 215 220  
 Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu  
 225 230 235

- 160 -

## (2) INFORMATION FOR SEQ ID NO:644:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 207 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:644:

```

Met Gly Gly Lys Trp Ser Lys Ser Ser Val Ile Gly Trp Pro Ala Val
 1           5           10           15
Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Gly Val Gly Ala
      20           25           30
Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr
      35           40           45
Ala Ala Asn Asn Ala Ala Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu
      50           55           60
Glu Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr
      65           70           75           80
Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
      85           90           95
Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
      100          105          110
Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
      115          120          125
Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
      130          135          140
Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu
      145          150          155          160
Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
      165          170          175
Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His
      180          185          190
His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys
      195          200          205

```

## (2) INFORMATION FOR SEQ ID NO:645:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:645:

```

Met Gly Gly Lys Ala Lys Ser Ser Val Val Arg His Ala Ser
 1           5           10           15
Gly Val Gly Ala Asn Asn Gln Cys
      20           25

```

- 161 -

## (2) INFORMATION FOR SEQ ID NO:646:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:646:

Met Gly Gly Lys Trp Ser Glu Ser Ser Val Val Arg Arg His Val Pro  
 1                      5                      10                      15  
 Leu Arg Gln Gly Ser Tyr Arg Ser Pro Leu  
                     20                      25

## (2) INFORMATION FOR SEQ ID NO:647:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 74 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:647:

CTACAAGGGA CTTTCCGCTG GGGACTTTCC AGGGAGGCGT GGCCTGGGCG GGACTGGGGA      60  
 GTGGCGAGCC CTCA      74

## (2) INFORMATION FOR SEQ ID NO:648:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 78 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:648:

CTGTTGGGGA CTTTCCATCC GTTGGGGACT TTCCAAGGCG GCGTGGCCTG GGTGACTAGT      60  
 TCCGGTGGGG ACTTTCCA      78

## (2) INFORMATION FOR SEQ ID NO:649:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 67 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:649:

CCGTTTGTTC CGTTGGGGAC TTTCCAGGAG ACGTGGCCTG AGTGACTAAG CCGCTGGGGA      60  
 CTTTCCG      67

- 162 -

## (2) INFORMATION FOR SEQ ID NO:650:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 621 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:650:

```

ATGGGTGGCA AGTGGTCAAA AAGTAGTGTG ATTGGATGGC CTGCTGTAAG GGAAAGAATG      60
AGACGAGCTG AGCCAGCAGC AGATGGGGTG GGAGCAGTAT CTCGAGACCT AGAAAAACAT      120
GGAGCAATCA CAAGTAGCAA TACAGCAGCT AACAAATGCTG CTTGTGCCTG GCTAGAAGCA      180
CAAGAGGAGG AAGAGGTGGG TTTTCCAGTC ACACCTCAGG TACCTTTAAG ACCAATGACT      240
TACAAGGCAG CTGTAGATCT TAGCCACTTT TTAAGAAGAA AGGGGGGACT GGAAGGGCTA      300
ATTCACTCCC AAAGAAGACA AGATATCCTT GATCTGTGGA TCTACCACAC ACAAGGCTAC      360
TTCCCTGATT GGCAGAACTA CACACCAGGG CCAGGGGTCA GATATCCACT GACCTTTGGA      420
TGGTGCTACA AGCTAGTACC AGTTGAGCCA GATAAGGTAG AAGAGGCCAA TAAAGGAGAG      480
AACACCAGCT TGTTACACCC TGTGAGCCTG CATGGAATGG ATGACCCTGA GAGAGAAGTG      540
TTAGAGTGGA GGTTCGACAG CCGCCTAGCA TTTCATCAGG TGGCCCGAGA GCTGCATCCG      600
GAGTACTTCA AGAACTGCTG A                                     621

```

## (2) INFORMATION FOR SEQ ID NO:651:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1596 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:651:

```

GAACAGATTT GGAATAACAT GACCTGGATG GAGTGGGACA GAGAAATTAA CAATTACACA      60
AGCTTAATAC ACTCCTTAAT TGAAGAATCG CAAAACCAGC AAGAAAAGAA TGAACAAGAA      120
TTATTGGAAT TAGATAAATG GGCAAGTTTG TGGAATTGGT TTAACATAAC AAATTGGCTG      180
TGGTATATAA AATTATTCAT AATGATAGTA GGAGGCTTGG TAGGTTTAAG AATAGTTTTT      240
GCTGTACTTT CTATAGTGAA TAGAGTTAGG CAGGGATATT CACCATTATC GTTTCAGACC      300
CACCTCCCAA TCCCGAGGGG ACCCGACAGG CCCGAAGGAA TAGAAGAAGA AGGTGGAGAG      360
AGAGACAGAG ACAGATCCAT TCGATTAGTG AACGGATCCT TAGCACTTAT CTGGGACGAT      420
CTGCGGAGCC TGTGCCTCTT CAGCTACCAC CGCTTGAGAG ACTTACTCTT GATTGTAACG      480
AGGATTGTGG AACTTCTGGG ACGCAGGGGG TGGGAAGCCC TCAAATATTG GTGGAATCTC      540
CTACAGTATT GGAGTCAGGA ACTAAGAAT AGTGCTGTTA ACTTGCTCAA TGCCACAGCC      600
ATAGCAGTAG CTGAGGGGAC AGATAGGGTT ATAGAAGTAT TACAAGCAGC TTATAGAGCT      660
ATTCGCCACA TACCTAGAAG AATAAGACAG GGCTTGAAA GGATTTTGCT ATAAGATGGG      720
TGGCAAGTGG TCAAAAAGTA GTGTGATTGG ATGGCCTGCT GTAAGGGAAA GAATGAGACG      780

```



- 163 -

AGCTGAGCCA GCAGCAGATG GGGTGGGAGC AGTATCTCGA GACCTAGAAA AACATGGAGC	840
AATCACAAGT AGCAATACAG CAGCTAACAA TGCTGCTTGT GCCTGGCTAG AAGCACAAGA	900
GGAGGAAGAG GTGGGTTTTT CAGTCACACC TCAGGTACCT TTAAGACCAA TGACTIONTCAA	960
GGCAGCTGTA GATCTTAGCC ACTTTTTTAAA AGAAAAGGGG GGACTIONGAAG GGCTAATTCA	1020
CTCCCAAAGA AGACAAGATA TCCTTGATCT GTGGATCTAC CACACACAAG GCTACTTCCC	1080
TGATTGGCAG AACTACACAC CAGGGCCAGG GGTGAGATAT CCACTIONGACCT TTGGATGGTG	1140
CTACAAGCTA GTACCAGTTG AGCCAGATAA GGTAGAAGAG GCCAATAAAG GAGAGAACAC	1200
CAGCTTGTTA CACCCTGTGA GCCTGCATGG AATGGATGAC CCTGAGAGAG AAGTGTTAGA	1260
GTGGAGGTTT GACAGCCGCC TAGCATTTC ACGCTGGCC CGAGAGCTGC ATCCGGAGTA	1320
CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG ACTTTCCAGG	1380
GAGGCGTGGC CTGGGCGGGA CTGGGGAGTG GCGAGCCCTC AGATGCTGCA TATAAGCAGC	1440
TGCTTTTTGC CTGTACTGGG TCTCTCTGGT TAGACCAGAT CTGAGCCTGG GAGCTCTCTG	1500
GCTAAGTAGG GAACCCACTG CTTAAGCCTC AATAAAGCTT GCCTTGAGTG CTTCAAGTAG	1560
TGTGTGCCCC TCTGTTGTGT GACTCTGGTA ACTAGA	1596

- 164 -

## (2) INFORMATION FOR SEQ ID NO:652:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:652:

GCTTTTGGCC

## (2) INFORMATION FOR SEQ ID NO:653:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:653:

CTTTTGGCCT

## (2) INFORMATION FOR SEQ ID NO:654:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:654:

TTTTTGGCCTG

## (2) INFORMATION FOR SEQ ID NO:655:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:655:

TTTTGGCCTGT

## (2) INFORMATION FOR SEQ ID NO:656:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:656:

TTTGCCTGTA

## (2) INFORMATION FOR SEQ ID NO:657:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:657:

TTGCCTGTAC

## (2) INFORMATION FOR SEQ ID NO:658:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:658:

TGCCTGTACT

## (2) INFORMATION FOR SEQ ID NO:659:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:659:

GCCTGTACTG

## (2) INFORMATION FOR SEQ ID NO:660:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:660:

CCTGTACTGG

- 165 -

## (2) INFORMATION FOR SEQ ID NO:661:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:661:

CTGTACTGGG

## (2) INFORMATION FOR SEQ ID NO:662:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:662:

TGTACTGGGT

## (2) INFORMATION FOR SEQ ID NO:663:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:663:

GTACTGGGTC

## (2) INFORMATION FOR SEQ ID NO:664:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:664:

TACTGGGTCT

## (2) INFORMATION FOR SEQ ID NO:665:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:665:

ACTGGGTCTC

## (2) INFORMATION FOR SEQ ID NO:666:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:666:

CTGGGTCTCT

## (2) INFORMATION FOR SEQ ID NO:667:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:667:

TGGGTCTCTC

## (2) INFORMATION FOR SEQ ID NO:668:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:668:

GGGTCTCTCT

## (2) INFORMATION FOR SEQ ID NO:669:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:669:

GGTCTCTCTG

## (2) INFORMATION FOR SEQ ID NO:670:

- 166 -

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:670:

GTCTCTCTGG

(2) INFORMATION FOR SEQ ID NO:671:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:671:

TCTCTCTGGT

(2) INFORMATION FOR SEQ ID NO:672:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:672:

CTCTCTGGTT

(2) INFORMATION FOR SEQ ID NO:673:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:673:

TCTCTGGTTA

(2) INFORMATION FOR SEQ ID NO:674:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:674:

CTCTGGTTAG

(2) INFORMATION FOR SEQ ID NO:675:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:675:

TCTCTGGTTA

(2) INFORMATION FOR SEQ ID NO:676:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:676:

CTGGTTAGAC

(2) INFORMATION FOR SEQ ID NO:677:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:677:

TGGTTAGACC

(2) INFORMATION FOR SEQ ID NO:678:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:678:

GGTTAGACCA

(2) INFORMATION FOR SEQ ID NO:679:

- (i) SEQUENCE CHARACTERISTICS:

- 167 -

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:679:

GTTAGACCAG

(2) INFORMATION FOR SEQ ID NO:680:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:680:

TTAGACCAGA

(2) INFORMATION FOR SEQ ID NO:681:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:681:

TAGACCAGAT

(2) INFORMATION FOR SEQ ID NO:682:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:682:

AGACCAGATC

(2) INFORMATION FOR SEQ ID NO:683:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:683:

GACCAGATCT

(2) INFORMATION FOR SEQ ID NO:684:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:684:

ACCAGATCTG

(2) INFORMATION FOR SEQ ID NO:685:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:685:

CCAGATCTGA

(2) INFORMATION FOR SEQ ID NO:686:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:686:

CAGATCTGAG

(2) INFORMATION FOR SEQ ID NO:687:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:687:

AGATCTGAGC

(2) INFORMATION FOR SEQ ID NO:688:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid

- 168 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:688:

GATCTGAGCC

(2) INFORMATION FOR SEQ ID NO:689:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:689:

ATCTGAGCCT

(2) INFORMATION FOR SEQ ID NO:690:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:690:

TCTGAGCCTG

(2) INFORMATION FOR SEQ ID NO:691:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:691:

CTGAGCCTGG

(2) INFORMATION FOR SEQ ID NO:692:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:692:

TGAGCCTGGG

(2) INFORMATION FOR SEQ ID NO:693:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:693:

GAGCCTGGGA

(2) INFORMATION FOR SEQ ID NO:694:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:694:

AGCCTGGGAG

(2) INFORMATION FOR SEQ ID NO:695:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:695:

GCCTGGGAGC

(2) INFORMATION FOR SEQ ID NO:696:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:696:

CCTGGGAGCT

(2) INFORMATION FOR SEQ ID NO:697:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 169 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:697:

CTGGGAGCTC

(2) INFORMATION FOR SEQ ID NO:698:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:698:

TGGGAGCTCT

(2) INFORMATION FOR SEQ ID NO:699:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:699:

GGGAGCTCTC

(2) INFORMATION FOR SEQ ID NO:700:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:700:

GGAGCTCTCT

(2) INFORMATION FOR SEQ ID NO:701:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:701:

GAGCTCTCTG

(2) INFORMATION FOR SEQ ID NO:702:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:702:

AGCTCTCTGG

(2) INFORMATION FOR SEQ ID NO:703:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:703:

GCTCTCTGGC

(2) INFORMATION FOR SEQ ID NO:704:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:704:

CTCTCTGGCT

(2) INFORMATION FOR SEQ ID NO:705:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:705:

TCTCTGGCTA

(2) INFORMATION FOR SEQ ID NO:706:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 170 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:706:

CTCTGGCTAA

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(2) INFORMATION FOR SEQ ID NO:707:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:707:

TCTGGCTAAC

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:711:

GCTAACTAGG

(2) INFORMATION FOR SEQ ID NO:712:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:712:

CTAACTAGGG

(2) INFORMATION FOR SEQ ID NO:708:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:708:

CTGGCTAACT

(2) INFORMATION FOR SEQ ID NO:713:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:713:

TAACTAGGGA

(2) INFORMATION FOR SEQ ID NO:709:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:709:

TGGCTAACTA

(2) INFORMATION FOR SEQ ID NO:714:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:714:

AACTAGGGAA

(2) INFORMATION FOR SEQ ID NO:710:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:710:

GGCTAACTAG

(2) INFORMATION FOR SEQ ID NO:715:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:715:

(2) INFORMATION FOR SEQ ID NO:711:

(i) SEQUENCE CHARACTERISTICS:



- 171 -

ACTAGGGAAC

## (2) INFORMATION FOR SEQ ID NO:716:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:716:

CTAGGGAACC

## (2) INFORMATION FOR SEQ ID NO:717:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:717:

TAGGGAACCC

## (2) INFORMATION FOR SEQ ID NO:718:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:718:

AGGGAACCCA

## (2) INFORMATION FOR SEQ ID NO:719:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:719:

GGGAACCCAC

## (2) INFORMATION FOR SEQ ID NO:720:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:720:

GGAACCCACT

## (2) INFORMATION FOR SEQ ID NO:721:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:721:

GAACCCACTG

## (2) INFORMATION FOR SEQ ID NO:722:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:722:

AACCCACTGC

## (2) INFORMATION FOR SEQ ID NO:723:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:723:

ACCCACTGCT

## (2) INFORMATION FOR SEQ ID NO:724:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:724:

CCCACTGCTT

- 172 -

## (2) INFORMATION FOR SEQ ID NO:725:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:725:

C CACTGCTTA

## (2) INFORMATION FOR SEQ ID NO:726:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:726:

C A C T G C T T A A

## (2) INFORMATION FOR SEQ ID NO:727:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:727:

A C T G C T T A A G

## (2) INFORMATION FOR SEQ ID NO:728:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:728:

C T G C T T A A G C

## (2) INFORMATION FOR SEQ ID NO:729:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:729:

T G C T T A A G C C

## (2) INFORMATION FOR SEQ ID NO:730:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:730:

G C T T A A G C C T

## (2) INFORMATION FOR SEQ ID NO:731:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:731:

C T T A A G C C T C

## (2) INFORMATION FOR SEQ ID NO:732:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:732:

T T A A G C C T C A

## (2) INFORMATION FOR SEQ ID NO:733:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:733:

T A A G C C T C A A

## (2) INFORMATION FOR SEQ ID NO:734:

- 173 -

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:734:

AAGCCTCAAT

(2) INFORMATION FOR SEQ ID NO:735:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:735:

AGCCTCAATA

(2) INFORMATION FOR SEQ ID NO:736:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:736:

GCCTCAATAA

(2) INFORMATION FOR SEQ ID NO:737:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:737:

CCTCAATAAA

(2) INFORMATION FOR SEQ ID NO:738:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:738:

CTCAATAAAG

(2) INFORMATION FOR SEQ ID NO:739:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:739:

TCAATAAAGC

(2) INFORMATION FOR SEQ ID NO:740:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:740:

CAATAAAGCT

(2) INFORMATION FOR SEQ ID NO:741:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:741:

AATAAAGCTT

(2) INFORMATION FOR SEQ ID NO:742:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:742:

ATAAAGCTTG

(2) INFORMATION FOR SEQ ID NO:743:

(i) SEQUENCE CHARACTERISTICS:

- 174 -

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:743:

TAAAGCTTGC

(2) INFORMATION FOR SEQ ID NO:744:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:744:

AAAGCTTGCC

(2) INFORMATION FOR SEQ ID NO:745:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:745:

AAGCTTGCCT

(2) INFORMATION FOR SEQ ID NO:746:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:746:

AGCTTGCCTT

(2) INFORMATION FOR SEQ ID NO:747:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:747:

GCTTGCCTTG

(2) INFORMATION FOR SEQ ID NO:748:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:748:

CTTGCCTTGA

(2) INFORMATION FOR SEQ ID NO:749:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:749:

TTGCCTTGAG

(2) INFORMATION FOR SEQ ID NO:750:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:750:

TGCCTTGAGT

(2) INFORMATION FOR SEQ ID NO:751:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:751:

GCCTTGAGTG

(2) INFORMATION FOR SEQ ID NO:752:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid

- 175 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:752:

CCTTGAGTGC

(2) INFORMATION FOR SEQ ID NO:753:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:753:

CTTGAGTGC

(2) INFORMATION FOR SEQ ID NO:754:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:754:

TTGAGTGCTT

(2) INFORMATION FOR SEQ ID NO:755:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:755:

TGAGTGCTTC

(2) INFORMATION FOR SEQ ID NO:756:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:756:

GAGTGCTTCA

(2) INFORMATION FOR SEQ ID NO:757:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:757:

AGTGCTTCAA

(2) INFORMATION FOR SEQ ID NO:758:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:758:

GTGCTTCAAG

(2) INFORMATION FOR SEQ ID NO:759:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:759:

TGCTTCAAGT

(2) INFORMATION FOR SEQ ID NO:760:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:760:

GCTTCAAGTA

(2) INFORMATION FOR SEQ ID NO:761:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 176 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:761:

CTTCAAGTAG

(2) INFORMATION FOR SEQ ID NO:762:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:762:

TTCAAGTAGT

(2) INFORMATION FOR SEQ ID NO:763:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:763:

TCAAGTAGTG

(2) INFORMATION FOR SEQ ID NO:764:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:764:

CAAGTAGTGT

(2) INFORMATION FOR SEQ ID NO:765:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:765:

AAGTAGTGTG

(2) INFORMATION FOR SEQ ID NO:766:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:766:

AGTAGTGTGT

(2) INFORMATION FOR SEQ ID NO:767:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:767:

GTAGTGTGTG

(2) INFORMATION FOR SEQ ID NO:768:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:768:

TAGTGTGTGC

(2) INFORMATION FOR SEQ ID NO:769:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:769:

AGTGTGTGCC

(2) INFORMATION FOR SEQ ID NO:770:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 177 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:770:

GTGTGTGCCC

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(2) INFORMATION FOR SEQ ID NO:771:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:771:

TGTGTGCCCCG

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:775:

TGCCCCGTCTG

(2) INFORMATION FOR SEQ ID NO:776:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:776:

GCCCCGTCTGT

(2) INFORMATION FOR SEQ ID NO:772:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:772:

GTGTGCCCCGT

(2) INFORMATION FOR SEQ ID NO:777:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:777:

CCCGTCTGTT

(2) INFORMATION FOR SEQ ID NO:773:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:773:

TGTGCCCCGTC

(2) INFORMATION FOR SEQ ID NO:778:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:778:

CCGTCTGTTG

(2) INFORMATION FOR SEQ ID NO:774:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:774:

GTGCCCCGTCT

(2) INFORMATION FOR SEQ ID NO:779:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:779:

(2) INFORMATION FOR SEQ ID NO:775:

(i) SEQUENCE CHARACTERISTICS:

- 178 -

CGTCTGTTGT

## (2) INFORMATION FOR SEQ ID NO:780:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:780:

GTCTGTTGTG

## (2) INFORMATION FOR SEQ ID NO:781:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:781:

TCTGTTGTGT

## (2) INFORMATION FOR SEQ ID NO:782:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:782:

CTGTTGTGTG

## (2) INFORMATION FOR SEQ ID NO:783:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:783:

TGTTGTGTGA

## (2) INFORMATION FOR SEQ ID NO:784:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:784:

GTTGTGTGAC

## (2) INFORMATION FOR SEQ ID NO:785:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:785:

TTGTGTGACT

## (2) INFORMATION FOR SEQ ID NO:786:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:786:

TGTTGTGACTC

## (2) INFORMATION FOR SEQ ID NO:787:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:787:

GTGTGACTCT

## (2) INFORMATION FOR SEQ ID NO:788:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:788:

TGTGTGACTC



- 179 -

## (2) INFORMATION FOR SEQ ID NO:789:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:789:

GTGTGACTCT

## (2) INFORMATION FOR SEQ ID NO:790:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:790:

TGTGACTCTG

## (2) INFORMATION FOR SEQ ID NO:791:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:791:

GTGACTCTGG

## (2) INFORMATION FOR SEQ ID NO:792:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:792:

TGACTCTGGT

## (2) INFORMATION FOR SEQ ID NO:793:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:793:

GACTCTGGTA

## (2) INFORMATION FOR SEQ ID NO:794:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:794:

ACTCTGGTAA

## (2) INFORMATION FOR SEQ ID NO:795:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:795:

CTCTGGTAAC

## (2) INFORMATION FOR SEQ ID NO:796:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:796:

TCTGGTAACT

## (2) INFORMATION FOR SEQ ID NO:797:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:797:

CTGGTAACTA

- 180 -

## (2) INFORMATION FOR SEQ ID NO:798:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:798:

TGGTAACTAG

## (2) INFORMATION FOR SEQ ID NO:799:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:799:

GGTAACTAGA

## (2) INFORMATION FOR SEQ ID NO:800:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 9207 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:800:

```
TGGAAGGGCT AATTCACCTCA CGGAAAAGAC CAGTTGAACC AGAAGAAGAT AGAAGAGGCC      60
ATGAAGAAGA AAACAACAGA TTGTTCTGCT TGCTCAGCTG GGGACTTTCC AGAAGGCGCG      120
GCCTGAGTGA CTAAGCCCCG TTGGGGACTT TCCGAAGAGG CATGAAGGGA CTTTCCAAGG      180
CAGGCGTGGC CTGGGCGGGA CTGGGGAGTG GCGAGCCCTC AGATGCTGCA TATAAGCAGC      240
TGCTTTCTGC CTGTACTGGG TCTCTCTGGT TAGACCAGAT CTGAGCCTGG GAGCTCTCTG      300
GCTAGCTAGG GAACCCACTG CTTAAGCCTC AATAAAGCTT GCCTTGAGTG CTTCAAGTAG      360
TGTGTGCCCC TCTGTTGTGT GACTCTGGTA TCTAGAGATC CCTCAGACCA TTTTAGTCCG      420
TGTGGAAAAT CTCTAGCAGT GGCGCCCGAA CAGGGGACTG AAAGCGAAAG GAAAACCAGA      480
GGAGCTCTCT CGACGCAGGA CTCGGCTTGC TGAAGCGCGC ACGGCAAGAG GCGAGGGGCG      540
GCGACTGGTG AGTACGCCGA AAATTTTGAC TAGCGGAGGC TAGAAGGAGA GAGATGGGTG      600
CGAGAGCGTC AATATTAAGC GGGGGAAAAT TAGATAGATG GGAGAAAATT CGGTTAAGGC      660
CAGGAGGAAA GAAAAAGTAT AAATTAAAC ATATAGTATG GGCAAGCAGG GAGCTAGAAC      720
GATTCGCAGT CAATCCTGGC CTGTTGGAAA CATCAGAAGG CTGTAGACAA ATACTGGGAC      780
AGTTACACCC GTCCCTTCAG ACAGGATCAG AAGAACTTAA ATCAGTATAT AATGCAGTAG      840
CAGTCCTCTA TTGTGTGCAT CAAAACATAG ACATAAAGGA CACCAAGGAA GCTTTAGAAA      900
AGATAGAGGA AGAGCAAAAC AAATGTAAGA AAAAAGCACA GCAAGCAGCA GCACAGCAAG      960
CAGCAGCTGG CACAGGAAAC AGCAACCCGG TCAGCCAAAA TTACCTTATA GTACAGAACA     1020
TGCAGGGGCA AATGGTACAT CAGGCCATAT CACCTAGAAC TTAAATGCA TGGGTAAAAG     1080
TAATAGAAGA GAAGGCTTTC AGCCCAGAGG TAATACCCAT GTTTTCAGCA TTATCAGAAG     1140
GAGCCACCCC ACAAGATTTA AACACCATGC TAAACACAGT GGGGGGACAT CAAGCAGCTA     1200
TGCAAATGTT AAAAGAGACC ATCAATGAGG AAGCTGCAGA ATGGGATAGA TTACATCCAG     1260
CGCAGGCAGG GCCTGTTGCA CCAGGCCAGA TGAGAGACCC AAGGGGAAGT GACATAGCAG     1320
GAACTACTAG TACCCTTCAG GAACAAATAG GATGGATGAC AGGTAATCCA GCTATCCCAG     1380
TAGGAGAAAT CTATAAAGA TGGATAATCC TGGGATTAAT TAAAATAGTA AGGATGTATA     1440
GCCCTATCAG CATTCTGGAC ATAAACAAG GACCAGGGA ACCCTTTAGA GACTATGTAG     1500
ACCGGTTCTA TAAACTCTA AGAGCCGAGC AAGCTACACA GGAGGTAAAA AATTGGATGA     1560
CAGAAACCTT GTTGGTCCAA AATGCAAACC CAGATTGTAA GACTATTTTA AAAGCATTGG     1620
```

GACCAGCAGC TACACTAGAA GAAATGATGA CAGCATGTCA GGGAGTGGGA GGACCCAGCC 1680  
ATAAAGCAAG AGTTTTGGCA GAAGCAATGA GCCAAGCAAC AAATGCAGCT ACTGTAATGA 1740  
TGCAGAGAAG CAATTTTAGA AACCAAAGAA AGAATGTTAA GTGTTTCAAT TGTGGCAAAG 1800  
AAGGGCACAT AGCCAGAAAT TGCAGGGCTC CTAGGAAAAG GGGCTGTTGG AAATGTGGAA 1860  
AGGAAGGACA CCAAATGAAA GATTGTACTG AGAGACAGGC TAATTTTTTA GGGAAAATCT 1920  
GGCCTTCCCA CAAGGGGAGG CCAGGGAAC TCTTCAGAG CAGGCCAGAA CCAACAGCCC 1980  
CTCTCCAGGG CAGGCCGGAG CCATCAGCCC CGCCAGAAGA GAGCTTCAGG TTTGGGGAGG 2040  
AGACAACAAC TCCCTCTCAG AAGCAGGAGC CGATAGACAG GGACAGGGAT CTGTATCCTT 2100  
TAGCTTCCCT CAGATCACTC TTTGGCAACG ACCCCTCGTC ACAATAAAGA TAGGGGGGCA 2160  
GCTGAAGGAA GCTCTATTAG ATACAGGAGC AGATGATACA GTATTAGAAG ACATGCATTT 2220  
GCCAGGAAAA TGGAAACCAA AAATGATAGG GGAATTGGA GGTTTTATCA AAGTAAACA 2280  
ATATGATGAA ATTCTTGTAG AAATCTGTGG ACATAAAGCT ATAGGTACAG TATTAGTAGG 2340  
ACCTACACCT GTCAACATAA TTGAAGAAA TCTGTTGACT CAGATTGGTT GCACTTTAAA 2400  
TTTTCCCAT AGTCCTATTG AAATGTACC AGTACAATTA AAGCCAGGAA TGGATGGCCC 2460  
AAAGGTTAAA CAATGGCCAT TGACAGAAGA GAAAATAAAA GCATTAGTAG AAATTTGTAC 2520  
AGAAATGGAA AAGGAAGGAA AGATTTCAA AATTGGGCCT GAAATCCAT ACAATACTCC 2580  
AGTATTTGCC ATAAAGAAAA AAGATGGTAC TAAATGGAGA AAATTAGTAG ATTTAGAGA 2640  
CCTTAATAAG AGAACTCAAG ACTTCTGGGA AGTTCAATTA GGAATACCAC ATCCCTCAGG 2700  
ATTAAAAAG AAAAAATCAG TAACAGTACT GGATGTGGGT GATGCATACT TTTCAGTTCC 2760  
CTTAGATGAA AACTTCAGGA AGTATACTGC ATTTACCATA CCTAGTATAA ATAATGAGAC 2820  
ACCAGGGATT AGATATCAGT ACAATGTGCT TCCACAGGGA TGGAAAGGAT CACCAGCAAT 2880  
ATTCCAAAGT AGCATGACAA GAATCTTAGA GCCTTTTAGA AGACAAATC CAGACATAGT 2940  
TATCTATCAA TACATGGATG ACTTGTATGT AGGATCTGAT TTAGAAATAG GACAGCATAG 3000  
AATAAAAATA GAGGAACTGA GACAACATCT GTTGAAGTGG GGATTACCA CACCAGACAA 3060  
AAAGCATCAG AAAGAACCCC CATTCTTTG GATGGGTTAT GAACTCCATC CTGATAAATG 3120  
GACAGTGCAA CCTATAGTAC TGCCAGAAAA AGACAGCTGG ACTGTCAATG ACATACAGAA 3180  
GTTAGTGGGT AAATTAAATT GGGCAAGTCA GATTTACCCA GGAATTAAAG TAAGGCAATT 3240  
ATGTAACTC CTTAGGGGAA CCAAGCACT AACAGAAGTA ATACCACTAA CAGAAGAAGC 3300  
AGAGCTAGAA CTGGCAGAAA ACAGGGAAAT TCTAAGAGAA CCAGTACATG GAGTGTATTA 3360  
TGACCCATCA AAAGACTTAA TAGCAGAAAT ACAGAAGCAG GAGCAAGGCC AATGGACATA 3420  
TCAAATTTAT CAAGATCAAT TTAATAATCT AAAACAGGA AAGTATGCAA GATTGAGGGG 3480  
TGCCCACT AATGATGTAA ACAAATTTCC AGAGGCAGTG CAAAAAATAG CCACAGAAAG 3540  
CATAGTAATA TGGGGAAAGA CTCCTAAATT TAGACTACCC ATACAAAAAG AAACATGGGA 3600  
CGCATGGTGG ACAGAGTATT GGCAAGCCAC CTGGATTCTT GAGTGGGAGT TTGTCAATAC 3660

- 183 -

CCCTCCCCTA GTAAAATTAT GGTACCAGTT AGAAAAAGAA CCCATAATAG GAGCAGAAAC	3720
TTTCTATGTA GATGGGGCAG CTAACAGAGA GACTAAATTA GGAAAAGCAG GATATGTTAC	3780
TGACAGAGGA AGACAAAAG TTGTCTCCCT AACTGACACA ACAAATCAGA AGACTGAGTT	3840
ACAAGCAATT CATCTAGCTT TGCAGGATTC AGGATTAGAA GTAAACATAG TAACAGACTC	3900
ACAGTATGCA TTAGGAATCA TTCAAGCACA ACCAGATAAA AGTGAATCAG AAATAGTCAA	3960
TCAAATAATA GAGCAATTAA TAAAAAGGA AAAGGTCTAC CTGGCATGGG TACCAGCACA	4020
CAAAGGAATT GGAGGGAATG AACAAGTAGA TAAATTAGTC AGTGCTGGAA TCAGGAAAAT	4080
ACTATTTTGA GATGGAATAG ATAAGGCACA AGAAGGCCAT GAGAAATATC ACAGTAATTG	4140
GAGAGCAATG GCTAGTGGTT TTAACCTGCC ACCTATAGTA GCAAAGAAA TAGTAGCCAG	4200
CTGTGATAAA TGTGAGCTAA AAGGAGAAGC CATGCATGGA CAAGTAGACT GTAGTCCAGG	4260
AATATGGCAA CTAGATTGTA CACATCTAGA AGGAAAAATT ATCCTGGTAG CAGTTCATGT	4320
AGCCAGTGGA TATATAGAAG CAGAAGTTAT TCCAGCAGAG ACAGGGCAGG AAACAGCATA	4380
CTTTATCTTA AAATTAGCAG GAAGGTGGCC AGTAAACACA ATACATACAG ACAATGGCGG	4440
CAATTTTCATC AGTACCACGG TTAAGGCCGC CTGTTGGTGG GCAGGGATCA AGCAGGAATT	4500
TGGCATTCCC TACAATCCCC AAAGCCAAGG AGTAGTGGA TCTATGAATA GAGAATTAAA	4560
GAAATTATA GGACAGGTAA GAGATCAGGC TGAACATCTT AAGACAGCAG TACAAATGGC	4620
AGTATTCATC CACAATTTTA AAAGAAAAGG GGGGATTGGG GGATACAGTG CAGGGGAAAG	4680
AATAGTAGAC ATAATAGCAA CAGACATACA AACTAAAGAA TTACAAAAGC AAATTACAAA	4740
AATTCAAAAT TTTCGGGTTT ATTACAGGGA CAGCAGAGAT CCACTTTGGA AAGGACCAGC	4800
AAAACCTCTC TGGAAAAGCG AAGGGGCAGT AGTAATACAA GATAATAGTG ACATAAAAGT	4860
AGTGCCAAGA AGAAAAGTAA AGATCATTAG GGATTATGGA AAACAGATGG CAGGTGATGA	4920
TTGTGTGGCA AGTAGACAGG ATGAGGATTA GAACATGGAA CAGTTTAGTG AAACACCATA	4980
TGTATGTTTC AAAGAAAGCT AAGGGATGGA TTTATAGACA TCACTATGAA AACACTCATC	5040
CAAAAATAAG CTCAGAAGTA CACATCCCAC TAGGGGAAGC TAGATTGGTA ATAACAACAT	5100
ATTGGGGTCT ACATACAGGA GAAAGAGACT GGCATTTGGG TCAGGGAGTC TCCATAGAAT	5160
GGAGGGAAAG GACATATAGA ACACAAGTAG ACCCCGAACT AGCAGACCAA CTAATTCATA	5220
TACATTACTT TGATTGTTTT TCAGAATCTG CCATAAGAAG TGCCATATTA GGATATAGAG	5280
TTAGGCATAG GTGTGAATAT CAAGCAGGAC ATAACAAGGT AGGATCTCTA CAGTACTTGG	5340
CACTAACAGC ATTAATAACA CCAAAGAAGA TAAAGCCACC TTTGCCTAGT GTTGCGAAAC	5400
TGACAGAGGA TAGATGGAAC AAGCCCCAGA AGACCAAGGG CCACAGAGGC AGCCATACAA	5460
TGAATGGACA CTAGAACTTT TAGAGGAGCT TAAGAATGAA GCTGTTAGGC ATTTTCCTAG	5520
GGTATGGCTC CATGGCTTAG GGCAACATAT CTATGAAACT TATGGGGATA CTTGGGAAGG	5580
AGTGGAGGCC ATAACAAGAA CTCTGCAACA ACTGCTGTTT ATTCATTTC AATTGGGTG	5640
TCAACATAGC AGAATAGGCA TTATTGACA GAGGAGAGCA AGAAATGGAG CCAGTAGATC	5700

CTAGACTAGA GCCCTGGAAG CATCCAGGAA GTCAGCCTAA GACTGCGTGT ACCACTTGCT 5760  
ATTGTAAAAA GTGCTGCTTT CATTGCCAAG TTTGTTTTAT GACAAAAGGC TTAGGCATCT 5820  
CCTATGGCAG GAAGAAGCGG AGACAGCGAC GAAGAGCTCC TCAAGACAGT CAGACTCATC 5880  
AAGCTTATCT ATCAAAGCAG TAAGTAATAT ATGTAATGCA ACCTTTACAA ATAGTAGCAA 5940  
TAGTAGCATT AGTAGTAGCA GGAATAATAG CAATAGTTGT GTGGACCATA GTATTCATAG 6000  
AATATAAGAA AATATTAAGA CAAAGAAAAA TAGACAGGTT GATTGATAGA ATAAGAGAAA 6060  
GAGCAGAAGA CAGTGGCAAT GACAGTGAAG GGGATCAGGA AGAATTATCG GCACTTGTGG 6120  
ACATGGGGCA CCATGATCCT TGGGATATTA ATGATCTGTA GAGCTGCAA CAATTTGTGG 6180  
GTCACAGTCT ATTATGGGGT ACCTGTGTGG AGAGAAGCAA CCACCACTCT ATTTTGTGCA 6240  
TCAGATGCCA AGGCATATGA TGCAGAGGTA CATAATGTTT GGGCCACACA TGCCTGTGTA 6300  
CCCACAGACC CTAACCCACA AGAAGTAGAA TTGAAAAATG TGACAGAAAA TTTTAACATG 6360  
TGGAAAAATA ACATGGTAGA ACAGATGCAT GAGGATATAA TCAGTTTATG GGATCAAAGC 6420  
CTGAAGCCAT GTGTAAAATT AACCCCACTC TGTGTTTCTT TAAATTGCAC TGATGCTACT 6480  
AATACCACTA ATAGTAATAC CACTAGCAGC AGCGAGAAAC CGAAGGGGAC AGGGGAAATA 6540  
AAAACTGCT CTTTCAATAT CACCACAAGC ATAAGAGATA AGGTGCAGAA ACAATATGCA 6600  
CTTTTTTATA GCCTTGATGT AGTACCAATG GATGATAATG ATAATAGTAC AAGCTATAGG 6660  
TTAATAAGTT GTAACACCTC AATCATTACA CAGGCCTGTC CAAAGATATC CTTTGAGCCA 6720  
ATTCCCATAC ATTATTGTGC CCCGGCTGGT TTTGCGATT CAAAGTGTA AGATAAAAGG 6780  
TTCAATGGAA AAGGACCATG TACAAGTGTC AGCACAGTAC AGTGACACA TGGGAATTAGG 6840  
CCAGTAGTAT CAACTCAACT GTTGTTAAAT GGCAGTCTAG CAGAAGAAGA GGTAGTAATT 6900  
AGATCTGACA ATTTTACGAA CAATGCTAAA ACCATAATAG TACAGCTGAG CAAATCTGTA 6960  
GAAATTACTT GTGTAAGACC CAACAACAAT ACAAGAAAAA GTATAAGTAT GGGACCAGGG 7020  
AGAGCATTTT ATACAACAGA AATAATAGGA GATATAAGAC AAGCATATTG TAACATTAGT 7080  
AAAGCAAAC TGGACTGACAC TTTAGAACAG ATAGCTAGAA AATTAAGAGA ACAATTTGAG 7140  
AATAAAACAA TAGTCTTTAA GCCATCCTCA GGAGGGGACC CAGAAATTGT AACACAGTTT 7200  
TACAGTTTTA ATTGTGGAGG GGAATTTTTC TACTGTAATT CAACACAACT GTTTAATGGT 7260  
ACTTGGAATG GTACTTGGGT TAATGGTACT TGGAGTAGTA ATAATACGAC TGATACTGCA 7320  
AATATCACAC TCCCATGCAG AATAAAACAA TTTATAAACA TGTGGCAGGA AGTAGGAAAA 7380  
GCAATGTATG CCCCTCCCAT CAAAGGACAA ATTAAATGTA CATCAAATAT TACAGGGCTG 7440  
ATATTAACAA GAGATGGTGG TAACAATAAC ACCACGAACG ACAACGAGAC CGAGACCTTC 7500  
AGACCTGGAG GAGGAGATAT GAGGGACAAT TGGAGAAGTG AATTATATAA ATATAAAGTA 7560  
GTACAAGTTG AACCATTAGG AGTAGCAGCC ACCAAGGCAA AGAGAAGAGT GGTGCAAAGA 7620  
GAAAAAGAG CAGTGGGAAT AGGAGCTATG TTCCTTGGGT TCTTAGGAGC AGCAGGAAGC 7680  
ACTATGGGCG CAGCGTCAGT GACGCTGACG GTACAAGCCA GACAATTATT GTCTGGTATA 7740

- 185 -

GTGCAGCAGC AGAACAATCT GCTGAGGGCT ATTGAGGCGC AACAGCATCT GTTGCAACTC	7800
ACAGTCTGGG GCATCAAACA GCTCCAGGCA AGAGTCCTGG CTGTGGAAAG ATACCTAAGG	7860
GATCAACAGC TCCTGGGACT TTGGGGTTGC TCTGGAAAAC TCATTGTCAC CACTACTGTG	7920
CCTTGGAACA ATAGCTGGAG TAATAAATCT CTGGAAACAA TTTGGGATAA CATGACCTGG	7980
ATGCAGTGGG AAAGAGAAAT TGACAATTAC ACAAACATAA TATACACCTT AATTGAAGAA	8040
TGCAGAACC AACAGAAAA AAATGAACTA GAATTATTGG AATTGGATAA ATGGGCAAAT	8100
TTGTGGAATT GGTTTAGTAT ATCAAATGG CTATGGTATA TAAAATTATT CATAATGGTA	8160
GTAGGAGGCT TGGTAGGTTT AAGAATAGTT TTTACTGTAC TTTCTATAGT TAATAGAGTT	8220
AGGCAGGGAT ACTCACCATT ATCGTTTCAG ACCCACCTCC CAACCCCGAA GGGACCCGAC	8280
AGGCCAGAAG GAATCGAAGA AGAAGGTGGA GAGAGAGACA GAGGCAGCTC CACTCGATTA	8340
GTGCACGGAT TCTTAGCACT TTTCTGGGAC GACCTGAGGA GTCTGTGCCT CTTGAGCTAC	8400
CACCACTTGA GAGACTTACT CTTGATTGTA ACGAGGATTG TGGAACCTCT GGGACGCAGG	8460
GGATGGGAAG CCCTCAAATA CTGGTGAAT CTCCTGCAGT ATTGGAGGCA GGAACCTACAG	8520
AAGAGTGCTG TTAGCTTGTT CAATGGCAGC GCCATAGCAG TAGCTGAGGG GACAGATAGA	8580
GTTATAGAAG CTTTACGAAG GGCTTATAGA GCTATTCTCC ACATACCTAG AAGAATAAGA	8640
CAGGGCTTAG AAAGGGCTTT GCTATAAAAT GGGTGGCAAG TGGTCAGAAA GTAGTGTGGT	8700
TAGAAGGCAT GTACCTTTAA GACAAGGCAG CTATAGATCT TAGCCGCTTT TAAAAGAAA	8760
AGGGGGGACT GGAAGGGCTA ATCACTCAC GGAAAAGACC AGTTGAACCA GAAGAAGATA	8820
GAAGAGGCCA TGAAGAAGAA AACAACAGAT TGTTCTGCTT GCTCAGCTGG GGACTTTCCA	8880
GAAGGCGCGG CCTGAGTGAC TAAGCCCCGT TGGGGACTTT CCGAAGAGGC ATGAAGGGAC	8940
TTTCCAAGGC AGGCGTGGCC TGGGCGGGAC TGGGGAGTGG CGAGCCCTCA GATGCTGCAT	9000
ATAAGCAGCT GCTTTCTGCC TGTACTGGGT CTCTCTGGTT AGACCAGATC TGAGCCTGGG	9060
AGCTCTCTGG CTAGCTAGGG AACCCACTGC TTAAGCCTCA ATAAAGCTTG CTTGAGTGC	9120
TTCAAGTAGT GTGTGCCCCGT CTGTTGTGTG ACTCTGGTAT CTAGAGATCC CTCAGACCAT	9180
TTTAGTCCGT GTGGAAAATC TCTAGCA	9207

## CLAIMS:

1. An isolated HIV-1 strain or a component, part, fragment or derivative thereof which is substantially non-pathogenic.
2. A strain of HIV-1 according to claim 1 wherein said strain is capable of stimulating an immune response in a human or primate against at least one glycoprotein on HIV-1 whilst not substantially reducing proliferative responses and cytokine production to a mitogen in said human or primate compared to a healthy, non-infected human or primate subject.
3. A strain of HIV-1 according to claim 2 which carries one or more mutations in its genome resulting in the inability to direct synthesis of at least one polypeptide or protein from a pathogenic strain of HIV-1 or results in reduced synthesis of said polypeptide or protein.
4. A strain of HIV-1 according to claim 2 which carries one or more mutations in its genome resulting in the said genome directing synthesis of a truncated form of a polypeptide or protein from a pathogenic strain of HIV-1.
5. A strain of HIV-1 according to claim 3 or 4 wherein said strain carries a mutation in its *nef* gene and/or long terminal repeat (LTR) region or in a functionally equivalent location in the HIV-1 genome.
6. A strain of HIV-1 according to claim 5 wherein said strain is reactive to antibodies to gp41-45, gp120 and/or gp160 of HIV-1 and carries a deletion of at least ten contiguous nucleotides in a region corresponding to the *nef* gene and/or LTR region of pathogenic HIV-1.
7. A strain of HIV-1 according to claim 6 wherein said strain is capable of inducing an immune response to at least one of *gag*, *pol* and/or *env*.



- 187 -

8. A strain of HIV-1 according to claim 7 wherein said strain carries a deletion in its genome of at least 10 nucleotides within the region from nucleotide 8787 to nucleotide 9709 using the nucleotide numbering of HIV-1 strain NL4-3.

9. A strain of HIV-1 according to claim 8 wherein said strain carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

nucleotide	(i)	8830-8862;
	(ii)	9009-9035;
	(iii)	9019-9029; and
	(iv)	9033-9049.

10. A strain of HIV-1 according to claim 8 wherein said strain carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

nucleotide	(v)	9281-9371;
	(vi)	9281-9362;
	(vii)	9105-9224; and
	(viii)	9271-9370.

11. A strain of HIV-1 according to claim 8 wherein said strain carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

nucleotide	(ix)	8882-8928;
	(x)	8850-9006;
	(xi)	8792-9041; and
	(xii)	9112-9204.

12. A strain of HIV-1 according to claim 8 wherein said strain carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

nucleotide	(xiii)	9105-9224
	(xiv)	9389-9395; and
	(xv)	9281-9366.

13. A strain of HIV-1 according to claim 8 wherein said deletion encompasses at least one of the following decanucleotides from the *nef* gene of HIV-1 strain NL4-3:

ATGGGTGGCA (SEQ ID NO:2);	TGGGTGGCAA (SEQ ID NO:3);
GGGTGGCAAG (SEQ ID NO:4);	GGTGGCAAGT (SEQ ID NO:5);
GTGGCAAGTG (SEQ ID NO:6);	TGGCAAGTGG (SEQ ID NO:7);
GGCAAGTGGT (SEQ ID NO:8);	GCAAGTGGTC (SEQ ID NO:9);
CAAGTGGTCA (SEQ ID NO:10);	AAGTGGTCAA (SEQ ID NO:11);
AGTGGTCAAA (SEQ ID NO:12);	GTGGTCAAAA (SEQ ID NO:13);
TGGTCAAAA (SEQ ID NO:14);	GGTCAAAAAG (SEQ ID NO:15);
GTCAAAAAGT (SEQ ID NO:16);	TCAAAAAGTA (SEQ ID NO:17);
CAAAAAGTAG (SEQ ID NO:18);	AAAAAGTAGT (SEQ ID NO:19);
AAAAGTAGTG (SEQ ID NO:20);	AAAGTAGTGT (SEQ ID NO:21);
AAGTAGTGTG (SEQ ID NO:22);	AGTAGTGTGA (SEQ ID NO:23);
GTAGTGTGAT (SEQ ID NO:24);	TAGTGTGATT (SEQ ID NO:25);
AGTGTGATTG (SEQ ID NO:26);	GTGTGATTGG (SEQ ID NO:27);
TGTGATTGGA (SEQ ID NO:28);	GTGATTGGAT (SEQ ID NO:29);
TGATTGGATG (SEQ ID NO:30);	GATTGGATGG (SEQ ID NO:31);
ATTGGATGGC (SEQ ID NO:32);	TTGGATGGCC (SEQ ID NO:33);
TGGATGGCCT (SEQ ID NO:34);	GGATGGCCTG (SEQ ID NO:35);
GATGGCCTGC (SEQ ID NO:36);	ATGGCCTGCT (SEQ ID NO:37);
TGGCCTGCTG (SEQ ID NO:38);	GGCCTGCTGT (SEQ ID NO:39);
GCCTGCTGTA (SEQ ID NO:40);	CCTGCTGTAA (SEQ ID NO:41);
CTGCTGTAAG (SEQ ID NO:42);	TGCTGTAAGG (SEQ ID NO:43);
GCTGTAAGGG (SEQ ID NO:44);	CTGTAAGGGA (SEQ ID NO:45);
TGTAAGGGAA (SEQ ID NO:46);	GTAAGGGAAA (SEQ ID NO:47);
TAAGGGAAAG (SEQ ID NO:48);	AAGGGAAAGA (SEQ ID NO:49);
AGGGAAAGAA (SEQ ID NO:50);	GGGAAAGAAT (SEQ ID NO:51);
GGAAAGAATG (SEQ ID NO:52);	GAAAGAATGA (SEQ ID NO:53);
AAAGAATGAG (SEQ ID NO:54);	AAGAATGAGA (SEQ ID NO:55);
AGAATGAGAC (SEQ ID NO:56);	GAATGAGACG (SEQ ID NO:57);
AATGAGACGA (SEQ ID NO:58);	ATGAGACGAG (SEQ ID NO:59);
TGAGACGAGC (SEQ ID NO:60);	GAGACGAGCT (SEQ ID NO:61);

- 189 -

AGACGAGCTG (SEQ ID NO:62);	GACGAGCTGA (SEQ ID NO:63);
ACGAGCTGAG (SEQ ID NO:64);	CGAGCTGAGC (SEQ ID NO:65);
GAGCTGAGCC (SEQ ID NO:66);	AGCTGAGCCA (SEQ ID NO:67);
GCTGAGCCAG (SEQ ID NO:68);	CTGAGCCAGC (SEQ ID NO:69);
TGAGCCAGCA (SEQ ID NO:70);	GAGCCAGCAG (SEQ ID NO:71);
AGCCAGCAGC (SEQ ID NO:72);	GCCAGCAGCA (SEQ ID NO:73);
CCAGCAGCAG (SEQ ID NO:74);	CAGCAGCAGA (SEQ ID NO:75);
AGCAGCAGAT (SEQ ID NO:76);	GCAGCAGATG (SEQ ID NO:77);
CAGCAGATGG (SEQ ID NO:78);	AGCAGATGGG (SEQ ID NO:79);
GCAGATGGGG (SEQ ID NO:80);	CAGATGGGGT (SEQ ID NO:81);
AGATGGGGTG (SEQ ID NO:82);	GATGGGGTGG (SEQ ID NO:83);
ATGGGGTGGG (SEQ ID NO:84);	TGGGGTGGGA (SEQ ID NO:85);
GGGGTGGGAG (SEQ ID NO:86);	GGGTGGGAGC (SEQ ID NO:87);
GGTGGGAGCA (SEQ ID NO:88);	GTGGGAGCAG (SEQ ID NO:89);
TGGGAGCAGT (SEQ ID NO:90);	GGGAGCAGTA (SEQ ID NO:91);
GGAGCAGTAT (SEQ ID NO:92);	GAGCAGTATC (SEQ ID NO:93);
AGCAGTATCT (SEQ ID NO:94);	GCAGTATCTC (SEQ ID NO:95);
CAGTATCTCG (SEQ ID NO:96);	AGTATCTCGA (SEQ ID NO:97);
GTATCTCGAG (SEQ ID NO:98);	TATCTCGAGA (SEQ ID NO:99);
ATCTCGAGAC (SEQ ID NO:100);	TCTCGAGACC (SEQ ID NO:101);
CTCGAGACCT (SEQ ID NO:102);	TCGAGACCTA (SEQ ID NO:103);
CGAGACCTAG (SEQ ID NO:104);	GAGACCTAGA (SEQ ID NO:105);
AGACCTAGAA (SEQ ID NO:106);	GACCTAGAAA (SEQ ID NO:107);
ACCTAGAAAA (SEQ ID NO:108);	CCTAGAAAAA (SEQ ID NO:109);
CTAGAAAAAC (SEQ ID NO:110);	TAGAAAAACA (SEQ ID NO:111);
AGAAAAACAT (SEQ ID NO:112);	GAAAAACATG (SEQ ID NO:113);
AAAAACATGG (SEQ ID NO:114);	AAAACATGGA (SEQ ID NO:115);
AAACATGGAG (SEQ ID NO:116);	AACATGGAGC (SEQ ID NO:117);
ACATGGAGCA (SEQ ID NO:118);	CATGGAGCAA (SEQ ID NO:119);
ATGGAGCAAT (SEQ ID NO:120);	TGGAGCAATC (SEQ ID NO:121);
GGAGCAATCA (SEQ ID NO:122);	GAGCAATCAC (SEQ ID NO:123);
AGCAATCACA (SEQ ID NO:124);	GCAATCACAA (SEQ ID NO:125);
CAATCACAAG (SEQ ID NO:126);	AATCACAAGT (SEQ ID NO:127);
ATCACAAGTA (SEQ ID NO:128);	TCACAAGTAG (SEQ ID NO:129);
CACAAGTAGC (SEQ ID NO:130);	ACAAGTAGCA (SEQ ID NO:131);
CAAGTAGCAA (SEQ ID NO:132);	AAGTAGCAAT (SEQ ID NO:133);
AGTAGCAATA (SEQ ID NO:134);	GTAGCAATAC (SEQ ID NO:135);
TAGCAATACA (SEQ ID NO:136);	AGCAATACAG (SEQ ID NO:137);

GCAATACAGC (SEQ ID NO:138);	CAATACAGCA (SEQ ID NO:139);
AATACAGCAG (SEQ ID NO:140);	ATACAGCAGC (SEQ ID NO:141);
TACAGCAGCT (SEQ ID NO:142);	ACAGCAGCTA (SEQ ID NO:143);
CAGCAGCTAA (SEQ ID NO:144);	AGCAGCTAAC (SEQ ID NO:145);
GCAGCTAACA (SEQ ID NO:146);	CAGCTAACAA (SEQ ID NO:147);
AGCTAACAAT (SEQ ID NO:148);	GCTAACAATG (SEQ ID NO:149);
CTAACAATGC (SEQ ID NO:150);	TAACAATGCT (SEQ ID NO:151);
AACAATGCTG (SEQ ID NO:152);	ACAATGCTGC (SEQ ID NO:153);
CAATGCTGCT (SEQ ID NO:154);	AATGCTGCTT (SEQ ID NO:155);
ATGCTGCTTG (SEQ ID NO:156);	TGCTGCTTGT (SEQ ID NO:157);
GCTGCTTGTG (SEQ ID NO:158);	CTGCTTGTGC (SEQ ID NO:159);
TGCTTGTGCC (SEQ ID NO:160);	GCTTGTGCCT (SEQ ID NO:161);
CTTGTGCCTG (SEQ ID NO:162);	TTGTGCCTGG (SEQ ID NO:163);
TGTGCCTGGC (SEQ ID NO:164);	GTGCCTGGCT (SEQ ID NO:165);
TGCCTGGCTA (SEQ ID NO:166);	GCCTGGCTAG (SEQ ID NO:167);
CCTGGCTAGA (SEQ ID NO:168);	CTGGCTAGAA (SEQ ID NO:169);
TGGCTAGAAG (SEQ ID NO:170);	GGCTAGAAGC (SEQ ID NO:171);
GCTAGAAGCA (SEQ ID NO:172);	CTAGAAGCAC (SEQ ID NO:173);
TAGAAGCACA (SEQ ID NO:174);	AGAAGCACAA (SEQ ID NO:175);
GAAGCACAAAG (SEQ ID NO:176);	AAGCACAAAG (SEQ ID NO:177);
AGCACAAAGAG (SEQ ID NO:178);	GCACAAGAGG (SEQ ID NO:179);
CACAAGAGGA (SEQ ID NO:180);	ACAAGAGGAG (SEQ ID NO:181);
CAAGAGGAGG (SEQ ID NO:182);	AAGAGGAGGA (SEQ ID NO:183);
AGAGGAGGAA (SEQ ID NO:184);	GAGGAGGAAG (SEQ ID NO:185);
AGGAGGAAGA (SEQ ID NO:186);	GGAGGAAGAG (SEQ ID NO:187);
GAGGAAGAGG (SEQ ID NO:188);	AGGAAGAGGT (SEQ ID NO:189);
GGAAGAGGTG (SEQ ID NO:190);	GAAGAGGTGG (SEQ ID NO:191);
AAGAGGTGGG (SEQ ID NO:192);	AGAGGTGGGT (SEQ ID NO:193);
GAGGTGGGTT (SEQ ID NO:194);	AGGTGGGTTT (SEQ ID NO:195);
GGTGGGTTTT (SEQ ID NO:196);	GTGGGTTTTT (SEQ ID NO:197);
TGGGTTTTTCC (SEQ ID NO:198);	GGGTTTTTCCA (SEQ ID NO:199);
GGTTTTTCCAG (SEQ ID NO:200);	GTTTTTCCAGT (SEQ ID NO:201);
TTTTTCCAGTC (SEQ ID NO:202);	TTTCCAGTCA (SEQ ID NO:203);
TTCCAGTCAC (SEQ ID NO:204);	TCCAGTCACA (SEQ ID NO:205);
CCAGTCACAC (SEQ ID NO:206);	CAGTCACACC (SEQ ID NO:207);
AGTCACACCT (SEQ ID NO:208);	GTCACACCTC (SEQ ID NO:209);
TCACACCTCA (SEQ ID NO:210);	CACACCTCAG (SEQ ID NO:211);
ACACCTCAGG (SEQ ID NO:212);	CACCTCAGGT (SEQ ID NO:213);

- 191 -

ACCTCAGGTA (SEQ ID NO:214);	CCTCAGGTAC (SEQ ID NO:215);
CTCAGGTACC (SEQ ID NO:216);	TCAGGTACCT (SEQ ID NO:217);
CAGGTACCTT (SEQ ID NO:218);	AGGTACCTTT (SEQ ID NO:219);
GGTACCTTTA (SEQ ID NO:220);	GTACCTTTAA (SEQ ID NO:221);
TACCTTTAAG (SEQ ID NO:222);	ACCTTTAAGA (SEQ ID NO:223);
CCTTTAAGAC (SEQ ID NO:224);	CTTTAAGACC (SEQ ID NO:225);
TTTAAGACCA (SEQ ID NO:226);	TTAAGACCAA (SEQ ID NO:227);
TAAGACCAAT (SEQ ID NO:228);	AAGACCAATG (SEQ ID NO:229);
AGACCAATGA (SEQ ID NO:230);	GACCAATGAC (SEQ ID NO:231);
ACCAATGACT (SEQ ID NO:232);	CCAATGACTT (SEQ ID NO:233);
CAATGACTTA (SEQ ID NO:234);	AATGACTTAC (SEQ ID NO:235);
ATGACTTACA (SEQ ID NO:236);	TGACTTACAA (SEQ ID NO:237);
GACTTACAAG (SEQ ID NO:238);	ACTTACAAGG (SEQ ID NO:239);
CTTACAAGGC (SEQ ID NO:240);	TTACAAGGCA (SEQ ID NO:241);
TACAAGGCAG (SEQ ID NO:242);	ACAAGGCAGC (SEQ ID NO:243);
CAAGGCAGCT (SEQ ID NO:244);	AAGGCAGCTG (SEQ ID NO:245);
AGGCAGCTGT (SEQ ID NO:246);	GGCAGCTGTA (SEQ ID NO:247);
GCAGCTGTAG (SEQ ID NO:248);	CAGCTGTAGA (SEQ ID NO:249);
AGCTGTAGAT (SEQ ID NO:250);	GCTGTAGATC (SEQ ID NO:251);
CTGTAGATCT (SEQ ID NO:252);	TGTAGATCTT (SEQ ID NO:253);
GTAGATCTTA (SEQ ID NO:254);	TAGATCTTAG (SEQ ID NO:255);
AGATCTTAGC (SEQ ID NO:256);	GATCTTAGCC (SEQ ID NO:257);
ATCTTAGCCA (SEQ ID NO:258);	TCTTAGCCAC (SEQ ID NO:259);
CTTAGCCACT (SEQ ID NO:260);	TTAGCCACTT (SEQ ID NO:261);
TAGCCACTTT (SEQ ID NO:262);	AGCCACTTTT (SEQ ID NO:263);
GCCACTTTTT (SEQ ID NO:264);	CCACTTTTTA (SEQ ID NO:265);
CACTTTTTAA (SEQ ID NO:266);	ACTTTTTTAA (SEQ ID NO:267);
CTTTTTAAAA (SEQ ID NO:268);	TTTTTTAAAAG (SEQ ID NO:269);
TTTTTAAAGA (SEQ ID NO:270);	TTTAAAGAA (SEQ ID NO:271);
TTAAAGAAA (SEQ ID NO:272);	TAAAGAAAA (SEQ ID NO:273);
AAAAGAAAAG (SEQ ID NO:274);	AAAGAAAAGG (SEQ ID NO:275);
AAGAAAAGGG (SEQ ID NO:276);	AGAAAAGGGG (SEQ ID NO:277);
GAAAAGGGGG (SEQ ID NO:278);	AAAAGGGGGG (SEQ ID NO:279);
AAAGGGGGGA (SEQ ID NO:280);	AAGGGGGGAC (SEQ ID NO:281);
AGGGGGGACT (SEQ ID NO:282);	GGGGGGACTG (SEQ ID NO:283);
GGGGGACTGG (SEQ ID NO:284);	GGGGACTGGA (SEQ ID NO:285);
GGGACTGGAA (SEQ ID NO:286);	GGACTGGAAG (SEQ ID NO:287);
GACTGGAAGG (SEQ ID NO:288);	ACTGGAAGGG (SEQ ID NO:289);

- 192 -

CTGGAAGGGC (SEQ ID NO:290);	TGGAAGGGCT (SEQ ID NO:291);
GGAAGGGCTA (SEQ ID NO:292);	GAAGGGCTAA (SEQ ID NO:293);
AAGGGCTAAT (SEQ ID NO:294);	AGGGCTAATT (SEQ ID NO:295);
GGGCTAATTC (SEQ ID NO:296);	GGCTAATTCA (SEQ ID NO:297);
GCTAATTCAC (SEQ ID NO:298);	CTAATTCACT (SEQ ID NO:299);
TAATTCACTC (SEQ ID NO:300);	AATTCACTCC (SEQ ID NO:301);
ATTCACTCCC (SEQ ID NO:302);	TTCACTCCCA (SEQ ID NO:303);
TCACTCCCAA (SEQ ID NO:304);	CACTCCCAA (SEQ ID NO:305);
ACTCCCAAAG (SEQ ID NO:306);	CTCCCAAAGA (SEQ ID NO:307);
TCCCAAAGAA (SEQ ID NO:308);	CCCAAAGAAG (SEQ ID NO:309);
CCAAAGAAGA (SEQ ID NO:310);	CAAAGAAGAC (SEQ ID NO:311);
AAAGAAGACA (SEQ ID NO:312);	AAGAAGACAA (SEQ ID NO:313);
AGAAGACAAG (SEQ ID NO:314);	GAAGACAAGA (SEQ ID NO:315);
AAGACAAGAT (SEQ ID NO:316);	AGACAAGATA (SEQ ID NO:317);
GACAAGATAT (SEQ ID NO:318);	ACAAGATATC (SEQ ID NO:319);
CAAGATATCC (SEQ ID NO:320);	AAGATATCCT (SEQ ID NO:321);
AGATATCCTT (SEQ ID NO:322);	GATATCCTTG (SEQ ID NO:323);
ATATCCTTGA (SEQ ID NO:324);	TATCCTTGAT (SEQ ID NO:325);
ATCCTTGATC (SEQ ID NO:326);	TCCTTGATCT (SEQ ID NO:327);
CCTTGATCTG (SEQ ID NO:328);	CTTGATCTGT (SEQ ID NO:329);
TTGATCTGTG (SEQ ID NO:330);	TGATCTGTGG (SEQ ID NO:331);
GATCTGTGGA (SEQ ID NO:332);	ATCTGTGGAT (SEQ ID NO:333);
TCTGTGGATC (SEQ ID NO:334);	CTGTGGATCT (SEQ ID NO:335);
TGTGGATCTA (SEQ ID NO:336);	GTGGATCTAC (SEQ ID NO:337);
TGGATCTACC (SEQ ID NO:338);	GGATCTACCA (SEQ ID NO:339);
GATCTACCAC (SEQ ID NO:340);	ATCTACCACA (SEQ ID NO:341);
TCTACCACAC (SEQ ID NO:342);	CTACCACACA (SEQ ID NO:343);
TACCACACAC (SEQ ID NO:344);	ACCACACACA (SEQ ID NO:345);
CCACACACAA (SEQ ID NO:346);	CACACACAAG (SEQ ID NO:347);
ACACACAAGG (SEQ ID NO:348);	CACACAAGGC (SEQ ID NO:349);
ACACAAGGCT (SEQ ID NO:350);	CACAAGGCTA (SEQ ID NO:351);
ACAAGGCTAC (SEQ ID NO:352);	CAAGGCTACT (SEQ ID NO:353);
AAGGCTACTT (SEQ ID NO:354);	AGGCTACTTC (SEQ ID NO:355);
GGCTACTTCC (SEQ ID NO:356);	GCTACTTCCC (SEQ ID NO:357);
CTACTTCCCT (SEQ ID NO:358);	TACTTCCCTG (SEQ ID NO:359);
ACTTCCCTGA (SEQ ID NO:360);	CTTCCCTGAT (SEQ ID NO:361);
TTCCCTGATT (SEQ ID NO:362);	TCCCTGATTG (SEQ ID NO:363);
CCCTGATTGG (SEQ ID NO:364);	CCTGATTGGC (SEQ ID NO:365);

CTGATTGGCA (SEQ ID NO:366);	TGATTGGCAG (SEQ ID NO:367);
GATTGGCAGA (SEQ ID NO:368);	ATTGGCAGAA (SEQ ID NO:369);
TTGGCAGAAC (SEQ ID NO:370);	TGGCAGAACT (SEQ ID NO:371);
GGCAGAACTA (SEQ ID NO:372);	GCAGAACTAC (SEQ ID NO:373);
CAGAACTACA (SEQ ID NO:374);	AGAACTACAC (SEQ ID NO:375);
GAACTACACA (SEQ ID NO:376);	AACTACACAC (SEQ ID NO:377);
ACTACACACC (SEQ ID NO:378);	CTACACACCA (SEQ ID NO:379);
TACACACCAG (SEQ ID NO:380);	ACACACCAGG (SEQ ID NO:381);
CACACCAGGG (SEQ ID NO:382);	ACACCAGGGC (SEQ ID NO:383);
CACCAGGGCC (SEQ ID NO:384);	ACCAGGGCCA (SEQ ID NO:385);
CCAGGGCCAG (SEQ ID NO:386);	CAGGGCCAGG (SEQ ID NO:387);
AGGGCCAGGG (SEQ ID NO:388);	GGGCCAGGGG (SEQ ID NO:389);
GGCCAGGGGT (SEQ ID NO:390);	GCCAGGGGTC (SEQ ID NO:391);
CCAGGGGTCA (SEQ ID NO:392);	CAGGGGTCAG (SEQ ID NO:393);
AGGGGTCAGA (SEQ ID NO:394);	GGGGTCAGAT (SEQ ID NO:395);
GGGTCAGATA (SEQ ID NO:396);	GGTCAGATAT (SEQ ID NO:397);
GTCAGATATC (SEQ ID NO:398);	TCAGATATCC (SEQ ID NO:399);
CAGATATCCA (SEQ ID NO:400);	AGATATCCAC (SEQ ID NO:401);
GATATCCACT (SEQ ID NO:402);	ATATCCACTG (SEQ ID NO:403);
TATCCACTGA (SEQ ID NO:404);	ATCCACTGAC (SEQ ID NO:405);
TCCACTGACC (SEQ ID NO:406);	CCACTGACCT (SEQ ID NO:407);
CACTGACCTT (SEQ ID NO:408);	ACTGACCTTT (SEQ ID NO:409);
CTGACCTTTG (SEQ ID NO:410);	TGACCTTTGG (SEQ ID NO:411);
GACCTTTGGA (SEQ ID NO:412);	ACCTTTGGAT (SEQ ID NO:413);
CCTTTGGATG (SEQ ID NO:414);	CTTTGGATGG (SEQ ID NO:415);
TTTGGATGGT (SEQ ID NO:416);	TTGGATGGTG (SEQ ID NO:417);
TGGATGGTGC (SEQ ID NO:418);	GGATGGTGCT (SEQ ID NO:419);
GATGGTGCTA (SEQ ID NO:420);	ATGGTGCTAC (SEQ ID NO:421);
TGGTGCTACA (SEQ ID NO:422);	GGTGCTACAA (SEQ ID NO:423);
GTGCTACAAG (SEQ ID NO:424);	TGCTACAAGC (SEQ ID NO:425);
GCTACAAGCT (SEQ ID NO:426);	CTACAAGCTA (SEQ ID NO:427);
TACAAGCTAG (SEQ ID NO:428);	ACAAGCTAGT (SEQ ID NO:429);
CAAGCTAGTA (SEQ ID NO:430);	AAGCTAGTAC (SEQ ID NO:431);
AGCTAGTACC (SEQ ID NO:432);	GCTAGTACCA (SEQ ID NO:433);
CTAGTACCAG (SEQ ID NO:434);	TAGTACCAGT (SEQ ID NO:435);
AGTACCAGTT (SEQ ID NO:436);	GTACCAGTTG (SEQ ID NO:437);
TACCAGTTGA (SEQ ID NO:438);	ACCAGTTGAG (SEQ ID NO:439);
CCAGTTGAGC (SEQ ID NO:440);	CAGTTGAGCC (SEQ ID NO:441);

- 194 -

AGTTGAGCCA (SEQ ID NO:442);	GTTGAGCCAG (SEQ ID NO:443);
TTGAGCCAGA (SEQ ID NO:444);	TGAGCCAGAT (SEQ ID NO:445);
GAGCCAGATA (SEQ ID NO:446);	AGCCAGATAA (SEQ ID NO:447);
GCCAGATAAG (SEQ ID NO:448);	CCAGATAAGG (SEQ ID NO:449);
CAGATAAGGT (SEQ ID NO:450);	AGATAAGGTA (SEQ ID NO:451);
GATAAGGTAG (SEQ ID NO:452);	ATAAGGTAGA (SEQ ID NO:453);
TAAGGTAGAA (SEQ ID NO:454);	AAGGTAGAAG (SEQ ID NO:455);
AGGTAGAAGA (SEQ ID NO:456);	GGTAGAAGAG (SEQ ID NO:457);
GTAGAAGAGG (SEQ ID NO:458);	TAGAAGAGGC (SEQ ID NO:459);
AGAAGAGGCC (SEQ ID NO:460);	GAAGAGGCCA (SEQ ID NO:461);
AAGAGGCCAA (SEQ ID NO:462);	AGAGGCCAAT (SEQ ID NO:463);
GAGGCCAATA (SEQ ID NO:464);	AGGCCAATAA (SEQ ID NO:465);
GGCCAATAAA (SEQ ID NO:466);	GCCAATAAAG (SEQ ID NO:467);
CCAATAAAGG (SEQ ID NO:468);	CAATAAAGGA (SEQ ID NO:469);
AATAAAGGAG (SEQ ID NO:470);	ATAAAGGAGA (SEQ ID NO:471);
TAAAGGAGAG (SEQ ID NO:472);	AAAGGAGAGA (SEQ ID NO:473);
AAGGAGAGAA (SEQ ID NO:474);	AGGAGAGAAC (SEQ ID NO:475);
GGAGAGAACA (SEQ ID NO:476);	GAGAGAACAC (SEQ ID NO:477);
AGAGAACACC (SEQ ID NO:478);	GAGAACACCA (SEQ ID NO:479);
AGAACACCAG (SEQ ID NO:480);	GAACACCAGC (SEQ ID NO:481);
AACACCAGCT (SEQ ID NO:482);	ACACCAGCTT (SEQ ID NO:483);
CACCAGCTTG (SEQ ID NO:484);	ACCAGCTTGT (SEQ ID NO:485);
CCAGCTTGTT (SEQ ID NO:486);	CAGCTTGTTA (SEQ ID NO:487);
AGCTTGTTAC (SEQ ID NO:488);	GCTTGTTACA (SEQ ID NO:489);
CTTGTTACAC (SEQ ID NO:490);	TTGTTACACC (SEQ ID NO:491);
TGTTACACCC (SEQ ID NO:492);	GTTACACCCT (SEQ ID NO:493);
TTACACCCTG (SEQ ID NO:494);	TACACCCTGT (SEQ ID NO:495);
ACACCCTGTG (SEQ ID NO:496);	CACCCTGTGA (SEQ ID NO:497);
ACCCTGTGAG (SEQ ID NO:498);	CCCTGTGAGC (SEQ ID NO:499);
CCTGTGAGCC (SEQ ID NO:500);	CTGTGAGCCT (SEQ ID NO:501);
TGTGAGCCTG (SEQ ID NO:502);	GTGAGCCTGC (SEQ ID NO:503);
TGAGCCTGCA (SEQ ID NO:504);	GAGCCTGCAT (SEQ ID NO:505);
AGCCTGCATG (SEQ ID NO:506);	GCCTGCATGG (SEQ ID NO:507);
CCTGCATGGA (SEQ ID NO:508);	CTGCATGGAA (SEQ ID NO:509);
TGCATGGAAT (SEQ ID NO:510);	GCATGGAATG (SEQ ID NO:511);
CATGGAATGG (SEQ ID NO:512);	ATGGAATGGA (SEQ ID NO:513);
TGGAATGGAT (SEQ ID NO:514);	GGAATGGATG (SEQ ID NO:515);
GAATGGATGA (SEQ ID NO:516);	AATGGATGAC (SEQ ID NO:517);



- 195 -

ATGGATGACC (SEQ ID NO:518);	TGGATGACCC (SEQ ID NO:519);
GGATGACCCT (SEQ ID NO:520);	GATGACCCTG (SEQ ID NO:521);
ATGACCCTGA (SEQ ID NO:522);	TGACCCTGAG (SEQ ID NO:523);
GACCCTGAGA (SEQ ID NO:524);	ACCCTGAGAG (SEQ ID NO:525);
CCCTGAGAGA (SEQ ID NO:526);	CCTGAGAGAG (SEQ ID NO:527);
CTGAGAGAGA (SEQ ID NO:528);	TGAGAGAGAA (SEQ ID NO:529);
GAGAGAGAAG (SEQ ID NO:530);	AGAGAGAAGT (SEQ ID NO:531);
GAGAGAAGTG (SEQ ID NO:532);	AGAGAAGTGT (SEQ ID NO:533);
GAGAAGTGTT (SEQ ID NO:534);	AGAAGTGTTA (SEQ ID NO:535);
GAAGTGTTAG (SEQ ID NO:536);	AAGTGTTAGA (SEQ ID NO:537);
AGTGTTAGAG (SEQ ID NO:538);	GTGTTAGAGT (SEQ ID NO:539);
TGTTAGAGTG (SEQ ID NO:540);	GTTAGAGTGG (SEQ ID NO:541);
TTAGAGTGGA (SEQ ID NO:542);	TAGAGTGAG (SEQ ID NO:543);
AGAGTGAGG (SEQ ID NO:544);	GAGTGAGGT (SEQ ID NO:545);
AGTGAGGTT (SEQ ID NO:546);	GTGAGGTTT (SEQ ID NO:547);
TGGAGGTTTG (SEQ ID NO:548);	GGAGGTTTGA (SEQ ID NO:549);
GAGGTTTGAC (SEQ ID NO:550);	AGGTTTGACA (SEQ ID NO:551);
GGTTTGACAG (SEQ ID NO:552);	GTTTGACAGC (SEQ ID NO:553);
TTTGACAGCC (SEQ ID NO:554);	TTGACAGCCG (SEQ ID NO:555);
TGACAGCCGC (SEQ ID NO:556);	GACAGCCGCC (SEQ ID NO:557);
ACAGCCGCCT (SEQ ID NO:558);	CAGCCGCCTA (SEQ ID NO:559);
AGCCGCCTAG (SEQ ID NO:560);	GCCGCCTAGC (SEQ ID NO:561);
CCGCCTAGCA (SEQ ID NO:562);	CGCCTAGCAT (SEQ ID NO:563);
GCCTAGCATT (SEQ ID NO:564);	CCTAGCATTT (SEQ ID NO:565);
CTAGCATTTT (SEQ ID NO:566);	TAGCATTTCA (SEQ ID NO:567);
AGCATTTTCAT (SEQ ID NO:568);	GCATTTTCATC (SEQ ID NO:569);
CATTTTCATCA (SEQ ID NO:570);	ATTTTCATCAC (SEQ ID NO:571);
TTTCATCACG (SEQ ID NO:572);	TTCATCACGT (SEQ ID NO:573);
TCATCACGTG (SEQ ID NO:574);	CATCACGTGG (SEQ ID NO:575);
ATCACGTGGC (SEQ ID NO:576);	TCACGTGGCC (SEQ ID NO:577);
CACGTGGCCC (SEQ ID NO:578);	ACGTGGCCCCG (SEQ ID NO:579);
CGTGGCCCCGA (SEQ ID NO:580);	GTGGCCCCGAG (SEQ ID NO:581);
TGGCCCCGAGA (SEQ ID NO:582);	GGCCCCGAGAG (SEQ ID NO:583);
GCCCCGAGAGC (SEQ ID NO:584);	CCCGAGAGCT (SEQ ID NO:585);
CCGAGAGCTG (SEQ ID NO:586);	CGAGAGCTGC (SEQ ID NO:587);
GAGAGCTGCA (SEQ ID NO:588);	AGAGCTGCAT (SEQ ID NO:589);
GAGCTGCATC (SEQ ID NO:590);	AGCTGCATCC (SEQ ID NO:591);
GCTGCATCCG (SEQ ID NO:592);	CTGCATCCGG (SEQ ID NO:593);

- 196 -

TGCATCCGGA (SEQ ID NO:594);	GCATCCGGAG (SEQ ID NO:595);
CATCCGGAGT (SEQ ID NO:596);	ATCCGGAGTA (SEQ ID NO:597);
TCCGGAGTAC (SEQ ID NO:598);	CCGGAGTACT (SEQ ID NO:599);
CGGAGTACTT (SEQ ID NO:600);	GGAGTACTTC (SEQ ID NO:601);
GAGTACTTCA (SEQ ID NO:602);	AGTACTTCAA (SEQ ID NO:603);
GTACTTCAAG (SEQ ID NO:604);	TACTTCAAGA (SEQ ID NO:605);
ACTTCAAGAA (SEQ ID NO:606);	CTTCAAGAAC (SEQ ID NO:607);
TTCAAGAACT (SEQ ID NO:608);	TCAAGAACTG (SEQ ID NO:609);
CAAGAACTGC (SEQ ID NO:610);	AAGAACTGCT (SEQ ID NO:611);
AGAACTGCTG (SEQ ID NO:612);	GAACTGCTGA (SEQ ID NO:613);

14. A strain of HIV-1 according to claim 8 wherein said deletion encompasses at least one of the following decanucleotides from the LTR region of HIV-1 strain NL4-3:

GCTTTTGGCC (SEQ ID NO:652);	CTTTTGCCT (SEQ ID NO:653);
TTTTTGCCTG (SEQ ID NO:654);	TTTTGCCTGT (SEQ ID NO:655);
TTTGCCTGTA (SEQ ID NO:656);	TTGCCTGTAC (SEQ ID NO:657);
TGCCTGTACT (SEQ ID NO:658);	GCCTGTACTG (SEQ ID NO:659);
CCTGTACTGG (SEQ ID NO:660);	CTGTACTGGG (SEQ ID NO:661);
TGTACTGGGT (SEQ ID NO:662);	GTACTGGGTC (SEQ ID NO:663);
TACTGGGTCT (SEQ ID NO:664);	ACTGGGTCTC (SEQ ID NO:665);
CTGGGTCTCT (SEQ ID NO:666);	TGGGTCTCTC (SEQ ID NO:667);
GGGTCTCTCT (SEQ ID NO:668);	GGTCTCTCTG (SEQ ID NO:669);
GTCTCTCTGG (SEQ ID NO:670);	TCTCTCTGGT (SEQ ID NO:671);
CTCTCTGGTT (SEQ ID NO:672);	TCTCTGGTTA (SEQ ID NO:673);
CTCTGGTTAG (SEQ ID NO:674);	TCTCTGGTTA (SEQ ID NO:675);
CTGGTTAGAC (SEQ ID NO:676);	TGGTTAGACC (SEQ ID NO:677);
GGTTAGACCA (SEQ ID NO:678);	GTTAGACCAG (SEQ ID NO:679);
TTAGACCAGA (SEQ ID NO:680);	TAGACCAGAT (SEQ ID NO:681);
AGACCAGATC (SEQ ID NO:682);	GACCAGATCT (SEQ ID NO:683);
ACCAGATCTG (SEQ ID NO:684);	CCAGATCTGA (SEQ ID NO:685);
CAGATCTGAG (SEQ ID NO:686);	AGATCTGAGC (SEQ ID NO:687);
GATCTGAGCC (SEQ ID NO:688);	ATCTGAGCCT (SEQ ID NO:689);
TCTGAGCCTG (SEQ ID NO:690);	CTGAGCCTGG (SEQ ID NO:691);
TGAGCCTGGG (SEQ ID NO:692);	GAGCCTGGGA (SEQ ID NO:693);
AGCCTGGGAG (SEQ ID NO:694);	GCCTGGGAGC (SEQ ID NO:695);
CCTGGGAGCT (SEQ ID NO:696);	CTGGGAGCTC (SEQ ID NO:697);
TGGGAGCTCT (SEQ ID NO:698);	GGGAGCTCTC (SEQ ID NO:699);

- 197 -

GGAGCTCTCT	(SEQ ID NO:700);	GAGCTCTCTG	(SEQ ID NO:701);
AGCTCTCTGG	(SEQ ID NO:702);	GCTCTCTGGC	(SEQ ID NO:703);
CTCTCTGGCT	(SEQ ID NO:704);	TCTCTGGCTA	(SEQ ID NO:705);
CTCTGGCTAA	(SEQ ID NO:706);	TCTGGCTAAC	(SEQ ID NO:707);
CTGGCTAACT	(SEQ ID NO:708);	TGGCTAACTA	(SEQ ID NO:709);
GGCTAACTAG	(SEQ ID NO:710);	GCTAACTAGG	(SEQ ID NO:711);
CTAACTAGGG	(SEQ ID NO:712);	TAACTAGGGA	(SEQ ID NO:713);
AACTAGGGAA	(SEQ ID NO:714);	ACTAGGGAAC	(SEQ ID NO:715);
CTAGGGAACC	(SEQ ID NO:716);	TAGGGAACCC	(SEQ ID NO:717);
AGGGAACCCA	(SEQ ID NO:718);	GGGAACCCAC	(SEQ ID NO:719);
GGAACCCACT	(SEQ ID NO:720);	GAACCCACTG	(SEQ ID NO:721);
AACCCACTGC	(SEQ ID NO:722);	ACCCACTGCT	(SEQ ID NO:723);
CCCACTGCTT	(SEQ ID NO:724);	CCACTGCTTA	(SEQ ID NO:725);
CACTGCTTAA	(SEQ ID NO:726);	ACTGCTTAAG	(SEQ ID NO:727);
CTGCTTAAGC	(SEQ ID NO:728);	TGCTTAAGCC	(SEQ ID NO:729);
GCTTAAGCCT	(SEQ ID NO:730);	CTTAAGCCTC	(SEQ ID NO:731);
TTAAGCCTCA	(SEQ ID NO:732);	TAAGCCTCAA	(SEQ ID NO:733);
AAGCCTCAAT	(SEQ ID NO:734);	AGCCTCAATA	(SEQ ID NO:735);
GCCTCAATAA	(SEQ ID NO:736);	CCTCAATAAA	(SEQ ID NO:737);
CTCAATAAAG	(SEQ ID NO:738);	TCAATAAAGC	(SEQ ID NO:739);
CAATAAAGCT	(SEQ ID NO:740);	AATAAAGCTT	(SEQ ID NO:741);
ATAAAGCTTG	(SEQ ID NO:742);	TAAAGCTTGC	(SEQ ID NO:743);
AAAGCTTGCC	(SEQ ID NO:744);	AAGCTTGCCCT	(SEQ ID NO:745);
AGCTTGCCCTT	(SEQ ID NO:746);	GCTTGCCCTTG	(SEQ ID NO:747);
CTTGCCCTGA	(SEQ ID NO:748);	TTGCCTTGAG	(SEQ ID NO:749);
TGCCTTGAGT	(SEQ ID NO:750);	GCCTTGAGTG	(SEQ ID NO:751);
CCTTGAGTGC	(SEQ ID NO:752);	CTTGAGTGCT	(SEQ ID NO:753);
TTGAGTGCTT	(SEQ ID NO:754);	TGAGTGCTTC	(SEQ ID NO:755);
GAGTGCTTCA	(SEQ ID NO:756);	AGTGCTTCAA	(SEQ ID NO:757);
GTGCTTCAAG	(SEQ ID NO:758);	TGCTTCAAGT	(SEQ ID NO:759);
GCTTCAAGTA	(SEQ ID NO:760);	CTTCAAGTAG	(SEQ ID NO:761);
TTCAAGTAGT	(SEQ ID NO:762);	TCAAGTAGTG	(SEQ ID NO:763);
CAAGTAGTGT	(SEQ ID NO:764);	AAGTAGTGTG	(SEQ ID NO:765);
AGTAGTGTGT	(SEQ ID NO:766);	GTAGTGTGTG	(SEQ ID NO:767);
TAGTGTGTGC	(SEQ ID NO:768);	AGTGTGTGCC	(SEQ ID NO:769);
GTGTGTGCCC	(SEQ ID NO:770);	TGTGTGCCCCG	(SEQ ID NO:771);
GTGTGCCCCGT	(SEQ ID NO:772);	TGTGCCCCGTC	(SEQ ID NO:773);
GTGCCCCGTCT	(SEQ ID NO:774);	TGCCCCGTCTG	(SEQ ID NO:775);

- 198 -

GCCCGTCTGT	(SEQ ID NO:776);	CCCGTCTGTT	(SEQ ID NO:777);
CCGTCTGTTG	(SEQ ID NO:778);	CGTCTGTTGT	(SEQ ID NO:779);
GTCTGTTGTG	(SEQ ID NO:780);	TCTGTTGTGT	(SEQ ID NO:781);
CTGTTGTGTG	(SEQ ID NO:782);	TGTTGTGTGA	(SEQ ID NO:783);
GTTGTGTGAC	(SEQ ID NO:784);	TTGTGTGACT	(SEQ ID NO:785);
TGTGTGACTC	(SEQ ID NO:786);	GTGTGACTCT	(SEQ ID NO:787);
TGTGTGACTC	(SEQ ID NO:788);	GTGTGACTCT	(SEQ ID NO:789);
TGTGACTCTG	(SEQ ID NO:790);	GTGACTCTGG	(SEQ ID NO:791);
TGACTCTGGT	(SEQ ID NO:792);	GACTCTGGTA	(SEQ ID NO:793);
ACTCTGGTAA	(SEQ ID NO:794);	CTCTGGTAAC	(SEQ ID NO:795);
TCTGGTAACT	(SEQ ID NO:796);	CTGGTAACTA	(SEQ ID NO:797);
TGGTAACTAG	(SEQ ID NO:798);	GGTAACTAGA	(SEQ ID NO:799).

15. A strain of HIV-1 according to claim 8 having the identifying characteristics of HIV-1 isolate ECACC V94101706 or HIV-1 isolate ECACC V941031169 or isolate C54 deposited at ECACC on 14 February, 1995 under Accession Number \_\_\_\_\_.

16. An isolated strain of HIV-1 which is reactive to antibodies to a glycoprotein of HIV-1, is capable of inducing an immune response to at least one of *gag*, *pol* and/or *env* and which is incapable of directing synthesis of a *nef* gene product or a full length *nef* gene product.

17. An isolated strain of HIV-1 according to claim 16 wherein the glycoprotein is at least one of gp41-45, gp120 and/or gp160.

18. A nucleic acid molecule or a part, fragment or derivative thereof from an HIV-1 isolate defined in any one of claims 1 to 16.

19. A molecular infectious clone comprising a nucleic acid molecule according to claim 18.

20. A method for inhibiting or reducing productive infection of an individual by a pathogenic strain of HIV-1, said method comprising administering to a subject a non-pathogenic isolate of HIV-1 in an amount effective to infect target cells and to generate

target cells carrying DNA derived from said non-pathogenic HIV-1.

21. A method according to claim 20 wherein said isolate is capable of stimulating an immune response in a human or primate against at least one glycoprotein on HIV-1 whilst not substantially reducing proliferative responses and cytokine production to a mitogen in said human or primate compared to a healthy, non-infected human or primate subject.

22. A method according to claim 21 wherein said isolate carries one or more mutations in its genome resulting in the inability to direct synthesis of at least one polypeptide or protein from a pathogenic strain of HIV-1.

23. A method according to claim 17 wherein said isolate carries one or more mutations in its genome resulting in said genome directing synthesis of a truncated form of a polypeptide or protein from a pathogenic strain of HIV-1.

24. A method according to claim 22 or 23 wherein said HIV-1 isolate carries a mutation in its *nef* gene and/or long terminal repeat (LTR) region or in a functionally equivalent location in the HIV-1 genome.

25. A method according to claim 24 wherein said HIV-1 isolate is reactive to antibodies to gp41-45, gp120 and/or gp160 of HIV-1 and carries a deletion of at least ten contiguous nucleotides in a region corresponding to the *nef* gene and/or LTR region of pathogenic HIV-1.

26. A method according to claim 25 wherein said HIV-1 isolate is capable of inducing an immune response to at least one of *gag*, *pol* and/or *env*.

27. A method according to claim 26 wherein said HIV-1 isolate carries a deletion in its genome of at least 10 nucleotides within the region from nucleotide 8787 to nucleotide 9709 using the nucleotide numbering of HIV-1 strain NL4-3.

- 200 -

28. A method according to claim 27 wherein said isolate carries a deletion in its genome of at least 10 nucleotides within a region selected from the list consisting of:

nucleotide	(i)	8830-8862;
	(ii)	9009-9035;
	(iii)	9019-9029; and
	(iv)	9033-9049.

29. A method according to claim 27 wherein said isolate carries a deletion in its genome of at least 10 nucleotides within a region selected from the list consisting of:

nucleotide	(v)	9281-9371;
	(vi)	9281-9362;
	(vii)	9105-9224; and
	(viii)	9271-9370.

30. A method according to claim 27 wherein said isolate carries a deletion in its genome of at least 10 nucleotides within a region selected from the list consisting of:

nucleotide	(ix)	8882-8928;
	(x)	8850-9006;
	(xi)	8792-9041; and
	(xii)	9112-9204.

31. A method according to claim 27 wherein said isolate carries a deletion in its genome of at least 10 nucleotides within a region selected from the list consisting of:

nucleotide	(xiii)	9105-9224;
	(xiv)	9389-9395; and
	(xv)	9281-9366.

32. A method according to claim 27 wherein said deletion encompasses at least one of the following decanucleotides from the nef gene of HIV-1 strain NL43:

ATGGGTGGCA (SEQ ID NO:2);	TGGGTGGCAA (SEQ ID NO:3);
GGGTGGCAAG (SEQ ID NO:4);	GGTGGCAAGT (SEQ ID NO:5);
GTGGCAAGTG (SEQ ID NO:6);	TGGCAAGTGG (SEQ ID NO:7);
GGCAAGTGGT (SEQ ID NO:8);	GCAAGTGGTC (SEQ ID NO:9);

- 201 -

CAAGTGGTCA (SEQ ID NO:10);	AAGTGGTCAA (SEQ ID NO:11);
AGTGGTCAAA (SEQ ID NO:12);	GTGGTCAAAA (SEQ ID NO:13);
TGGTCAAAAA (SEQ ID NO:14);	GGTCAAAAAG (SEQ ID NO:15);
GTCAAAAAGT (SEQ ID NO:16);	TCAAAAAGTA (SEQ ID NO:17);
CAAAAAGTAG (SEQ ID NO:18);	AAAAAGTAGT (SEQ ID NO:19);
AAAAGTAGTG (SEQ ID NO:20);	AAAGTAGTGT (SEQ ID NO:21);
AAGTAGTGTG (SEQ ID NO:22);	AGTAGTGTGA (SEQ ID NO:23);
GTAGTGTGAT (SEQ ID NO:24);	TAGTGTGATT (SEQ ID NO:25);
AGTGTGATTG (SEQ ID NO:26);	GTGTGATTGG (SEQ ID NO:27);
TGTGATTGGA (SEQ ID NO:28);	GTGATTGGAT (SEQ ID NO:29);
TGATTGGATG (SEQ ID NO:30);	GATTGGATGG (SEQ ID NO:31);
ATTGGATGGC (SEQ ID NO:32);	TTGGATGGCC (SEQ ID NO:33);
TGGATGGCCT (SEQ ID NO:34);	GGATGGCCTG (SEQ ID NO:35);
GATGGCCTGC (SEQ ID NO:36);	ATGGCCTGCT (SEQ ID NO:37);
TGGCCTGCTG (SEQ ID NO:38);	GGCCTGCTGT (SEQ ID NO:39);
GCCTGCTGTA (SEQ ID NO:40);	CCTGCTGTAA (SEQ ID NO:41);
CTGCTGTAAG (SEQ ID NO:42);	TGCTGTAAGG (SEQ ID NO:43);
GCTGTAAGGG (SEQ ID NO:44);	CTGTAAGGGA (SEQ ID NO:45);
TGTAAGGGAA (SEQ ID NO:46);	GTAAGGGAAA (SEQ ID NO:47);
TAAGGGAAAG (SEQ ID NO:48);	AAGGGAAAGA (SEQ ID NO:49);
AGGGAAAGAA (SEQ ID NO:50);	GGGAAAGAAT (SEQ ID NO:51);
GGAAAGAATG (SEQ ID NO:52);	GAAAGAATGA (SEQ ID NO:53);
AAAGAATGAG (SEQ ID NO:54);	AAGAATGAGA (SEQ ID NO:55);
AGAATGAGAC (SEQ ID NO:56);	GAATGAGACG (SEQ ID NO:57);
AATGAGACGA (SEQ ID NO:58);	ATGAGACGAG (SEQ ID NO:59);
TGAGACGAGC (SEQ ID NO:60);	GAGACGAGCT (SEQ ID NO:61);
AGACGAGCTG (SEQ ID NO:62);	GACGAGCTGA (SEQ ID NO:63);
ACGAGCTGAG (SEQ ID NO:64);	CGAGCTGAGC (SEQ ID NO:65);
GAGCTGAGCC (SEQ ID NO:66);	AGCTGAGCCA (SEQ ID NO:67);
GCTGAGCCAG (SEQ ID NO:68);	CTGAGCCAGC (SEQ ID NO:69);
TGAGCCAGCA (SEQ ID NO:70);	GAGCCAGCAG (SEQ ID NO:71);
AGCCAGCAGC (SEQ ID NO:72);	GCCAGCAGCA (SEQ ID NO:73);
CCAGCAGCAG (SEQ ID NO:74);	CAGCAGCAGA (SEQ ID NO:75);
AGCAGCAGAT (SEQ ID NO:76);	GCAGCAGATG (SEQ ID NO:77);
CAGCAGATGG (SEQ ID NO:78);	AGCAGATGGG (SEQ ID NO:79);
GCAGATGGGG (SEQ ID NO:80);	CAGATGGGGT (SEQ ID NO:81);
AGATGGGGTG (SEQ ID NO:82);	GATGGGGTGG (SEQ ID NO:83);
ATGGGGTGGG (SEQ ID NO:84);	TGGGGTGGGA (SEQ ID NO:85);

- 202 -

GGGGTGGGAG (SEQ ID NO:86);	GGGTGGGAGC (SEQ ID NO:87);
GGTGGGAGCA (SEQ ID NO:88);	GTGGGAGCAG (SEQ ID NO:89);
TGGGAGCAGT (SEQ ID NO:90);	GGGAGCAGTA (SEQ ID NO:91);
GGAGCAGTAT (SEQ ID NO:92);	GAGCAGTATC (SEQ ID NO:93);
AGCAGTATCT (SEQ ID NO:94);	GCAGTATCTC (SEQ ID NO:95);
CAGTATCTCG (SEQ ID NO:96);	AGTATCTCGA (SEQ ID NO:97);
GTATCTCGAG (SEQ ID NO:98);	TATCTCGAGA (SEQ ID NO:99);
ATCTCGAGAC (SEQ ID NO:100);	TCTCGAGACC (SEQ ID NO:101);
CTCGAGACCT (SEQ ID NO:102);	TCGAGACCTA (SEQ ID NO:103);
CGAGACCTAG (SEQ ID NO:104);	GAGACCTAGA (SEQ ID NO:105);
AGACCTAGAA (SEQ ID NO:106);	GACCTAGAAA (SEQ ID NO:107);
ACCTAGAAAA (SEQ ID NO:108);	CCTAGAAAAA (SEQ ID NO:109);
CTAGAAAAAC (SEQ ID NO:110);	TAGAAAAACA (SEQ ID NO:111);
AGAAAAACAT (SEQ ID NO:112);	GAAAAACATG (SEQ ID NO:113);
AAAAACATGG (SEQ ID NO:114);	AAAACATGGA (SEQ ID NO:115);
AAACATGGAG (SEQ ID NO:116);	AACATGGAGC (SEQ ID NO:117);
ACATGGAGCA (SEQ ID NO:118);	CATGGAGCAA (SEQ ID NO:119);
ATGGAGCAAT (SEQ ID NO:120);	TGGAGCAATC (SEQ ID NO:121);
GGAGCAATCA (SEQ ID NO:122);	GAGCAATCAC (SEQ ID NO:123);
AGCAATCACA (SEQ ID NO:124);	GCAATCACAA (SEQ ID NO:125);
CAATCACAAG (SEQ ID NO:126);	AATCACAAGT (SEQ ID NO:127);
ATCACAAGTA (SEQ ID NO:128);	TCACAAGTAG (SEQ ID NO:129);
CACAAGTAGC (SEQ ID NO:130);	ACAAGTAGCA (SEQ ID NO:131);
CAAGTAGCAA (SEQ ID NO:132);	AAGTAGCAAT (SEQ ID NO:133);
AGTAGCAATA (SEQ ID NO:134);	GTAGCAATAC (SEQ ID NO:135);
TAGCAATACA (SEQ ID NO:136);	AGCAATACAG (SEQ ID NO:137);
GCAATACAGC (SEQ ID NO:138);	CAATACAGCA (SEQ ID NO:139);
AATACAGCAG (SEQ ID NO:140);	ATACAGCAGC (SEQ ID NO:141);
TACAGCAGCT (SEQ ID NO:142);	ACAGCAGCTA (SEQ ID NO:143);
CAGCAGCTAA (SEQ ID NO:144);	AGCAGCTAAC (SEQ ID NO:145);
GCAGCTAACA (SEQ ID NO:146);	CAGCTAACAA (SEQ ID NO:147);
AGCTAACAAT (SEQ ID NO:148);	GCTAACAATG (SEQ ID NO:149);
CTAACAATGC (SEQ ID NO:150);	TAACAATGCT (SEQ ID NO:151);
AACAATGCTG (SEQ ID NO:152);	ACAATGCTGC (SEQ ID NO:153);
CAATGCTGCT (SEQ ID NO:154);	AATGCTGCTT (SEQ ID NO:155);
ATGCTGCTTG (SEQ ID NO:156);	TGCTGCTTGT (SEQ ID NO:157);
GCTGCTTGTG (SEQ ID NO:158);	CTGCTTGTGC (SEQ ID NO:159);
TGCTTGTGCC (SEQ ID NO:160);	GCTTGTGCCT (SEQ ID NO:161);



- 203 -

CTTGTGCCTG (SEQ ID NO:162);	TTGTGCCTGG (SEQ ID NO:163);
TGTGCCTGGC (SEQ ID NO:164);	GTGCCTGGCT (SEQ ID NO:165);
TGCCTGGCTA (SEQ ID NO:166);	GCCTGGCTAG (SEQ ID NO:167);
CCTGGCTAGA (SEQ ID NO:168);	CTGGCTAGAA (SEQ ID NO:169);
TGGCTAGAAG (SEQ ID NO:170);	GGCTAGAAGC (SEQ ID NO:171);
GCTAGAAGCA (SEQ ID NO:172);	CTAGAAGCAC (SEQ ID NO:173);
TAGAAGCACA (SEQ ID NO:174);	AGAAGCACAA (SEQ ID NO:175);
GAAGCACAAG (SEQ ID NO:176);	AAGCACAAGA (SEQ ID NO:177);
AGCACAAGAG (SEQ ID NO:178);	GCACAAGAGG (SEQ ID NO:179);
CACAAGAGGA (SEQ ID NO:180);	ACAAGAGGAG (SEQ ID NO:181);
CAAGAGGAGG (SEQ ID NO:182);	AAGAGGAGGA (SEQ ID NO:183);
AGAGGAGGAA (SEQ ID NO:184);	GAGGAGGAAG (SEQ ID NO:185);
AGGAGGAAGA (SEQ ID NO:186);	GGAGGAAGAG (SEQ ID NO:187);
GAGGAAGAGG (SEQ ID NO:188);	AGGAAGAGGT (SEQ ID NO:189);
GGAAGAGGTG (SEQ ID NO:190);	GAAGAGGTGG (SEQ ID NO:191);
AAGAGGTGGG (SEQ ID NO:192);	AGAGGTGGGT (SEQ ID NO:193);
GAGGTGGGTT (SEQ ID NO:194);	AGGTGGGTTT (SEQ ID NO:195);
GGTGGGTTTT (SEQ ID NO:196);	GTGGGTTTTT (SEQ ID NO:197);
TGGGTTTTTCC (SEQ ID NO:198);	GGGTTTTTCCA (SEQ ID NO:199);
GGTTTTTCCAG (SEQ ID NO:200);	GTTTTTCCAGT (SEQ ID NO:201);
TTTTTCCAGTC (SEQ ID NO:202);	TTTCCAGTCA (SEQ ID NO:203);
TTCCAGTCAC (SEQ ID NO:204);	TCCAGTCACA (SEQ ID NO:205);
CCAGTCACAC (SEQ ID NO:206);	CAGTCACACC (SEQ ID NO:207);
AGTCACACCT (SEQ ID NO:208);	GTCACACCTC (SEQ ID NO:209);
TCACACCTCA (SEQ ID NO:210);	CACACCTCAG (SEQ ID NO:211);
ACACCTCAGG (SEQ ID NO:212);	CACCTCAGGT (SEQ ID NO:213);
ACCTCAGGTA (SEQ ID NO:214);	CCTCAGGTAC (SEQ ID NO:215);
CTCAGGTACC (SEQ ID NO:216);	TCAGGTACCT (SEQ ID NO:217);
CAGGTACCTT (SEQ ID NO:218);	AGGTACCTTT (SEQ ID NO:219);
GGTACCTTTA (SEQ ID NO:220);	GTACCTTTAA (SEQ ID NO:221);
TACCTTTAAG (SEQ ID NO:222);	ACCTTTAAGA (SEQ ID NO:223);
CCTTTAAGAC (SEQ ID NO:224);	CTTTAAGACC (SEQ ID NO:225);
TTTAAGACCA (SEQ ID NO:226);	TTAAGACCAA (SEQ ID NO:227);
TAAGACCAAT (SEQ ID NO:228);	AAGACCAATG (SEQ ID NO:229);
AGACCAATGA (SEQ ID NO:230);	GACCAATGAC (SEQ ID NO:231);
ACCAATGACT (SEQ ID NO:232);	CCAATGACTT (SEQ ID NO:233);
CAATGACTTA (SEQ ID NO:234);	AATGACTTAC (SEQ ID NO:235);
ATGACTTACA (SEQ ID NO:236);	TGACTTACAA (SEQ ID NO:237);

- 204 -

GACTTACAAG (SEQ ID NO:238);	ACTTACAAGG (SEQ ID NO:239);
CTTACAAGGC (SEQ ID NO:240);	TTACAAGGCA (SEQ ID NO:241);
TACAAGGCAG (SEQ ID NO:242);	ACAAGGCAGC (SEQ ID NO:243);
CAAGGCAGCT (SEQ ID NO:244);	AAGGCAGCTG (SEQ ID NO:245);
AGGCAGCTGT (SEQ ID NO:246);	GGCAGCTGTA (SEQ ID NO:247);
GCAGCTGTAG (SEQ ID NO:248);	CAGCTGTAGA (SEQ ID NO:249);
AGCTGTAGAT (SEQ ID NO:250);	GCTGTAGATC (SEQ ID NO:251);
CTGTAGATCT (SEQ ID NO:252);	TGTAGATCTT (SEQ ID NO:253);
GTAGATCTTA (SEQ ID NO:254);	TAGATCTTAG (SEQ ID NO:255);
AGATCTTAGC (SEQ ID NO:256);	GATCTTAGCC (SEQ ID NO:257);
ATCTTAGCCA (SEQ ID NO:258);	TCTTAGCCAC (SEQ ID NO:259);
CTTAGCCACT (SEQ ID NO:260);	TTAGCCACTT (SEQ ID NO:261);
TAGCCACTTT (SEQ ID NO:262);	AGCCACTTTT (SEQ ID NO:263);
GCCACTTTTT (SEQ ID NO:264);	CCACTTTTTA (SEQ ID NO:265);
CACTTTTTAA (SEQ ID NO:266);	ACTTTTTAAA (SEQ ID NO:267);
CTTTTTAAAA (SEQ ID NO:268);	TTTTTAAAG (SEQ ID NO:269);
TTTTTAAAGA (SEQ ID NO:270);	TTTTTAAAGAA (SEQ ID NO:271);
TTAAAGAAA (SEQ ID NO:272);	TAAAGAAAA (SEQ ID NO:273);
AAAAGAAAAG (SEQ ID NO:274);	AAAGAAAAGG (SEQ ID NO:275);
AAGAAAAGGG (SEQ ID NO:276);	AGAAAAGGGG (SEQ ID NO:277);
GAAAAGGGGG (SEQ ID NO:278);	AAAAGGGGGG (SEQ ID NO:279);
AAAGGGGGGA (SEQ ID NO:280);	AAGGGGGGAC (SEQ ID NO:281);
AGGGGGGACT (SEQ ID NO:282);	GGGGGGACTG (SEQ ID NO:283);
GGGGGACTGG (SEQ ID NO:284);	GGGGACTGGA (SEQ ID NO:285);
GGGACTGGA (SEQ ID NO:286);	GGACTGGAAG (SEQ ID NO:287);
GACTGGAAGG (SEQ ID NO:288);	ACTGGAAGGG (SEQ ID NO:289);
CTGGAAGGGC (SEQ ID NO:290);	TGGAAGGGCT (SEQ ID NO:291);
GGAAGGGCTA (SEQ ID NO:292);	GAAGGGCTAA (SEQ ID NO:293);
AAGGGCTAAT (SEQ ID NO:294);	AGGGCTAATT (SEQ ID NO:295);
GGGCTAATTC (SEQ ID NO:296);	GGCTAATTCA (SEQ ID NO:297);
GCTAATTCAC (SEQ ID NO:298);	CTAATTCACT (SEQ ID NO:299);
TAATTCACTC (SEQ ID NO:300);	AATTCACTCC (SEQ ID NO:301);
ATTCACTCCC (SEQ ID NO:302);	TTCCTCCCA (SEQ ID NO:303);
TCACTCCCAA (SEQ ID NO:304);	CACTCCCAAA (SEQ ID NO:305);
ACTCCCAAAG (SEQ ID NO:306);	CTCCCAAAGA (SEQ ID NO:307);
TCCCAAAGAA (SEQ ID NO:308);	CCCAAAGAAG (SEQ ID NO:309);
CCAAAGAAGA (SEQ ID NO:310);	CAAAGAAGAC (SEQ ID NO:311);
AAAGAAGACA (SEQ ID NO:312);	AAGAAGACAA (SEQ ID NO:313);

- 205 -

AGAAGACAAG (SEQ ID NO:314);	GAAGACAAGA (SEQ ID NO:315);
AAGACAAGAT (SEQ ID NO:316);	AGACAAGATA (SEQ ID NO:317);
GACAAGATAT (SEQ ID NO:318);	ACAAGATATC (SEQ ID NO:319);
CAAGATATCC (SEQ ID NO:320);	AAGATATCCT (SEQ ID NO:321);
AGATATCCTT (SEQ ID NO:322);	GATATCCTTG (SEQ ID NO:323);
ATATCCTTGA (SEQ ID NO:324);	TATCCTTGAT (SEQ ID NO:325);
ATCCTTGATC (SEQ ID NO:326);	TCCTTGATCT (SEQ ID NO:327);
CCTTGATCTG (SEQ ID NO:328);	CTTGATCTGT (SEQ ID NO:329);
TTGATCTGTG (SEQ ID NO:330);	TGATCTGTGG (SEQ ID NO:331);
GATCTGTGGA (SEQ ID NO:332);	ATCTGTGGAT (SEQ ID NO:333);
TCTGTGGATC (SEQ ID NO:334);	CTGTGGATCT (SEQ ID NO:335);
TGTGGATCTA (SEQ ID NO:336);	GTGGATCTAC (SEQ ID NO:337);
TGGATCTACC (SEQ ID NO:338);	GGATCTACCA (SEQ ID NO:339);
GATCTACCAC (SEQ ID NO:340);	ATCTACCACA (SEQ ID NO:341);
TCTACCACAC (SEQ ID NO:342);	CTACCACACA (SEQ ID NO:343);
TACCACACAC (SEQ ID NO:344);	ACCACACACA (SEQ ID NO:345);
CCACACACAA (SEQ ID NO:346);	CACACACAAG (SEQ ID NO:347);
ACACACAAGG (SEQ ID NO:348);	CACACAAGGC (SEQ ID NO:349);
ACACAAGGCT (SEQ ID NO:350);	CACAAGGCTA (SEQ ID NO:351);
ACAAGGCTAC (SEQ ID NO:352);	CAAGGCTACT (SEQ ID NO:353);
AAGGCTACTT (SEQ ID NO:354);	AGGCTACTTC (SEQ ID NO:355);
GGCTACTTCC (SEQ ID NO:356);	GCTACTTCCC (SEQ ID NO:357);
CTACTTCCCT (SEQ ID NO:358);	TACTTCCCTG (SEQ ID NO:359);
ACTTCCCTGA (SEQ ID NO:360);	CTTCCCTGAT (SEQ ID NO:361);
TTCCCTGATT (SEQ ID NO:362);	TCCCTGATTG (SEQ ID NO:363);
CCCTGATTGG (SEQ ID NO:364);	CCTGATTGGC (SEQ ID NO:365);
CTGATTGGCA (SEQ ID NO:366);	TGATTGGCAG (SEQ ID NO:367);
GATTGGCAGA (SEQ ID NO:368);	ATTGGCAGAA (SEQ ID NO:369);
TTGGCAGAAC (SEQ ID NO:370);	TGGCAGAACT (SEQ ID NO:371);
GGCAGAACTA (SEQ ID NO:372);	GCAGAACTAC (SEQ ID NO:373);
CAGAACTACA (SEQ ID NO:374);	AGAACTACAC (SEQ ID NO:375);
GAACTACACA (SEQ ID NO:376);	AACTACACAC (SEQ ID NO:377);
ACTACACACC (SEQ ID NO:378);	CTACACACCA (SEQ ID NO:379);
TACACACCAG (SEQ ID NO:380);	ACACACCAGG (SEQ ID NO:381);
CACACCAGGG (SEQ ID NO:382);	ACACCAGGGC (SEQ ID NO:383);
CACCAGGGCC (SEQ ID NO:384);	ACCAGGGCCA (SEQ ID NO:385);
CCAGGGCCAG (SEQ ID NO:386);	CAGGGCCAGG (SEQ ID NO:387);
AGGGCCAGGG (SEQ ID NO:388);	GGGCCAGGGG (SEQ ID NO:389);

- 206 -

GGCCAGGGGT (SEQ ID NO:390);	GCCAGGGGTC (SEQ ID NO:391);
CCAGGGGTCA (SEQ ID NO:392);	CAGGGGTCAG (SEQ ID NO:393);
AGGGGTCAGA (SEQ ID NO:394);	GGGGTCAGAT (SEQ ID NO:395);
GGGTCAGATA (SEQ ID NO:396);	GGTCAGATAT (SEQ ID NO:397);
GTCAGATATC (SEQ ID NO:398);	TCAGATATCC (SEQ ID NO:399);
CAGATATCCA (SEQ ID NO:400);	AGATATCCAC (SEQ ID NO:401);
GATATCCACT (SEQ ID NO:402);	ATATCCACTG (SEQ ID NO:403);
TATCCACTGA (SEQ ID NO:404);	ATCCACTGAC (SEQ ID NO:405);
TCCACTGACC (SEQ ID NO:406);	CCACTGACCT (SEQ ID NO:407);
CACTGACCTT (SEQ ID NO:408);	ACTGACCTTT (SEQ ID NO:409);
CTGACCTTTG (SEQ ID NO:410);	TGACCTTTGG (SEQ ID NO:411);
GACCTTTGGA (SEQ ID NO:412);	ACCTTTGGAT (SEQ ID NO:413);
CCTTTGGATG (SEQ ID NO:414);	CTTTGGATGG (SEQ ID NO:415);
TTTGGATGGT (SEQ ID NO:416);	TTGGATGGTG (SEQ ID NO:417);
TGGATGGTGCT (SEQ ID NO:418);	GGATGGTGCT (SEQ ID NO:419);
GATGGTGCTA (SEQ ID NO:420);	ATGGTGCTAC (SEQ ID NO:421);
TGGTGCTACA (SEQ ID NO:422);	GGTGCTACAA (SEQ ID NO:423);
GTGCTACAAG (SEQ ID NO:424);	TGCTACAAGC (SEQ ID NO:425);
GCTACAAGCT (SEQ ID NO:426);	CTACAAGCTA (SEQ ID NO:427);
TACAAGCTAG (SEQ ID NO:428);	ACAAGCTAGT (SEQ ID NO:429);
CAAGCTAGTA (SEQ ID NO:430);	AAGCTAGTAC (SEQ ID NO:431);
AGCTAGTACC (SEQ ID NO:432);	GCTAGTACCA (SEQ ID NO:433);
CTAGTACCAG (SEQ ID NO:434);	TAGTACCAGT (SEQ ID NO:435);
AGTACCAGTT (SEQ ID NO:436);	GTACCAGTTG (SEQ ID NO:437);
TACCAGTTGA (SEQ ID NO:438);	ACCAGTTGAG (SEQ ID NO:439);
CCAGTTGAGC (SEQ ID NO:440);	CAGTTGAGCC (SEQ ID NO:441);
AGTTGAGCCA (SEQ ID NO:442);	GTTGAGCCAG (SEQ ID NO:443);
TTGAGCCAGA (SEQ ID NO:444);	TGAGCCAGAT (SEQ ID NO:445);
GAGCCAGATA (SEQ ID NO:446);	AGCCAGATAA (SEQ ID NO:447);
GCCAGATAAG (SEQ ID NO:448);	CCAGATAAGG (SEQ ID NO:449);
CAGATAAGGT (SEQ ID NO:450);	AGATAAGGTA (SEQ ID NO:451);
GATAAGGTAG (SEQ ID NO:452);	ATAAGGTAGA (SEQ ID NO:453);
TAAGGTAGAA (SEQ ID NO:454);	AAGGTAGAAG (SEQ ID NO:455);
AGGTAGAAGA (SEQ ID NO:456);	GGTAGAAGAG (SEQ ID NO:457);
GTAGAAGAGG (SEQ ID NO:458);	TAGAAGAGGC (SEQ ID NO:459);
AGAAGAGGCC (SEQ ID NO:460);	GAAGAGGCCA (SEQ ID NO:461);
AAGAGGCCAA (SEQ ID NO:462);	AGAGGCCAAT (SEQ ID NO:463);
GAGGCCAATA (SEQ ID NO:464);	AGGCCAATAA (SEQ ID NO:465);

- 207 -

GGCCAATAAA (SEQ ID NO:466);	GCCAATAAAG (SEQ ID NO:467);
CCAATAAAGG (SEQ ID NO:468);	CAATAAAGGA (SEQ ID NO:469);
AATAAAGGAG (SEQ ID NO:470);	ATAAAGGAGA (SEQ ID NO:471);
TAAAGGAGAG (SEQ ID NO:472);	AAAGGAGAGA (SEQ ID NO:473);
AAGGAGAGAA (SEQ ID NO:474);	AGGAGAGAAC (SEQ ID NO:475);
GGAGAGAAAC (SEQ ID NO:476);	GAGAGAACAC (SEQ ID NO:477);
AGAGAACACC (SEQ ID NO:478);	GAGAACACCA (SEQ ID NO:479);
AGAACACCAG (SEQ ID NO:480);	GAACACCAGC (SEQ ID NO:481);
AACACCAGCT (SEQ ID NO:482);	ACACCAGCTT (SEQ ID NO:483);
CACCAGCTTG (SEQ ID NO:484);	ACCAGCTTGT (SEQ ID NO:485);
CCAGCTTGTT (SEQ ID NO:486);	CAGCTTGTTA (SEQ ID NO:487);
AGCTTGTTAC (SEQ ID NO:488);	GCTTGTTACA (SEQ ID NO:489);
CTTGTTACAC (SEQ ID NO:490);	TTTGTTACACC (SEQ ID NO:491);
TGTACACCC (SEQ ID NO:492);	GTTACACCCT (SEQ ID NO:493);
TTACACCCTG (SEQ ID NO:494);	TACACCCTGT (SEQ ID NO:495);
ACACCCTGTG (SEQ ID NO:496);	CACCCTGTGA (SEQ ID NO:497);
ACCCTGTGAG (SEQ ID NO:498);	CCCTGTGAGC (SEQ ID NO:499);
CCTGTGAGCC (SEQ ID NO:500);	CTGTGAGCCT (SEQ ID NO:501);
TGTGAGCCTG (SEQ ID NO:502);	GTGAGCCTGC (SEQ ID NO:503);
TGAGCCTGCA (SEQ ID NO:504);	GAGCCTGCAT (SEQ ID NO:505);
AGCCTGCATG (SEQ ID NO:506);	GCCTGCATGG (SEQ ID NO:507);
CCTGCATGGA (SEQ ID NO:508);	CTGCATGGAA (SEQ ID NO:509);
TGCATGGAAT (SEQ ID NO:510);	GCATGGAATG (SEQ ID NO:511);
CATGGAATGG (SEQ ID NO:512);	ATGGAATGGA (SEQ ID NO:513);
TGGAATGGAT (SEQ ID NO:514);	GGAATGGATG (SEQ ID NO:515);
GAATGGATGA (SEQ ID NO:516);	AATGGATGAC (SEQ ID NO:517);
ATGGATGACC (SEQ ID NO:518);	TGGATGACCC (SEQ ID NO:519);
GGATGACCCT (SEQ ID NO:520);	GATGACCCTG (SEQ ID NO:521);
ATGACCCTGA (SEQ ID NO:522);	TGACCCTGAG (SEQ ID NO:523);
GACCCTGAGA (SEQ ID NO:524);	ACCCTGAGAG (SEQ ID NO:525);
CCCTGAGAGA (SEQ ID NO:526);	CCTGAGAGAG (SEQ ID NO:527);
CTGAGAGAGA (SEQ ID NO:528);	TGAGAGAGAA (SEQ ID NO:529);
GAGAGAGAAG (SEQ ID NO:530);	AGAGAGAAGT (SEQ ID NO:531);
GAGAGAAGTG (SEQ ID NO:532);	AGAGAAGTGT (SEQ ID NO:533);
GAGAAGTGTT (SEQ ID NO:534);	AGAAGTGTTA (SEQ ID NO:535);
GAAGTGTTAG (SEQ ID NO:536);	AAGTGTTAGA (SEQ ID NO:537);
AGTGTTAGAG (SEQ ID NO:538);	GTGTTAGAGT (SEQ ID NO:539);
TGTTAGAGTG (SEQ ID NO:540);	GTTAGAGTGG (SEQ ID NO:541);

TTAGAGTGGA (SEQ ID NO:542);	TAGAGTGGAG (SEQ ID NO:543);
AGAGTGGAGG (SEQ ID NO:544);	GAGTGGAGGT (SEQ ID NO:545);
AGTGGAGGTT (SEQ ID NO:546);	GTGGAGGTTT (SEQ ID NO:547);
TGGAGGTTTG (SEQ ID NO:548);	GGAGGTTTGA (SEQ ID NO:549);
GAGGTTTGAC (SEQ ID NO:550);	AGGTTTGACA (SEQ ID NO:551);
GGTTTGACAG (SEQ ID NO:552);	GTTTGACAGC (SEQ ID NO:553);
TTTGACAGCC (SEQ ID NO:554);	TTGACAGCCG (SEQ ID NO:555);
TGACAGCCGC (SEQ ID NO:556);	GACAGCCGCC (SEQ ID NO:557);
ACAGCCGCCT (SEQ ID NO:558);	CAGCCGCCTA (SEQ ID NO:559);
AGCCGCCTAG (SEQ ID NO:560);	GCCGCCTAGC (SEQ ID NO:561);
CCGCCTAGCA (SEQ ID NO:562);	CGCCTAGCAT (SEQ ID NO:563);
GCCTAGCATT (SEQ ID NO:564);	CCTAGCATTT (SEQ ID NO:565);
CTAGCATTTT (SEQ ID NO:566);	TAGCATTTCA (SEQ ID NO:567);
AGCATTTTCAT (SEQ ID NO:568);	GCATTTTCATC (SEQ ID NO:569);
CATTTTCATCA (SEQ ID NO:570);	ATTTTCATCAC (SEQ ID NO:571);
TTTCATCACG (SEQ ID NO:572);	TTCATCACGT (SEQ ID NO:573);
TCATCACGTG (SEQ ID NO:574);	CATCACGTGG (SEQ ID NO:575);
ATCACGTGGC (SEQ ID NO:576);	TCACGTGGCC (SEQ ID NO:577);
CACGTGGCCC (SEQ ID NO:578);	ACGTGGCCCC (SEQ ID NO:579);
CGTGGCCCCG (SEQ ID NO:580);	GTGGCCCCGAG (SEQ ID NO:581);
TGGCCCCGAGA (SEQ ID NO:582);	GGCCCCGAGAG (SEQ ID NO:583);
GCCCCGAGAGC (SEQ ID NO:584);	CCCGAGAGCT (SEQ ID NO:585);
CCGAGAGCTG (SEQ ID NO:586);	CGAGAGCTGC (SEQ ID NO:587);
GAGAGCTGCA (SEQ ID NO:588);	AGAGCTGCAT (SEQ ID NO:589);
GAGCTGCATC (SEQ ID NO:590);	AGCTGCATCC (SEQ ID NO:591);
GCTGCATCCG (SEQ ID NO:592);	CTGCATCCGG (SEQ ID NO:593);
TGCATCCGGA (SEQ ID NO:594);	GCATCCGGAG (SEQ ID NO:595);
CATCCGGAGT (SEQ ID NO:596);	ATCCGGAGTA (SEQ ID NO:597);
TCCGGAGTAC (SEQ ID NO:598);	CCGGAGTACT (SEQ ID NO:599);
CGGAGTACTT (SEQ ID NO:600);	GGAGTACTTC (SEQ ID NO:601);
GAGTACTTCA (SEQ ID NO:602);	AGTACTTCAA (SEQ ID NO:603);
GTA TTTCAAG (SEQ ID NO:604);	TACTTCAAGA (SEQ ID NO:605);
ACTTCAAGAA (SEQ ID NO:606);	CTTCAAGAAC (SEQ ID NO:607);
TTCAAGAACT (SEQ ID NO:608);	TCAAGAACTG (SEQ ID NO:609);
CAAGAACTGC (SEQ ID NO:610);	AAGAACTGCT (SEQ ID NO:611);
AGAACTGCTG (SEQ ID NO:612);	GAACTGCTGA (SEQ ID NO:613);

33. A method according to claim 27 wherein said deletion encompasses at least one of the following decanucleotides from the LTR region of HIV-1 strain NL4-3:

GCTTTTTGCC	(SEQ ID NO:652);	CTTTTTGCCT	(SEQ ID NO:653);
TTTTTGCCTG	(SEQ ID NO:654);	TTTTGCCTGT	(SEQ ID NO:655);
TTTGCCTGTA	(SEQ ID NO:656);	TTGCCTGTAC	(SEQ ID NO:657);
TGCCTGTACT	(SEQ ID NO:658);	GCCTGTACTG	(SEQ ID NO:659);
CCTGTACTGG	(SEQ ID NO:660);	CTGTACTGGG	(SEQ ID NO:661);
TGTACTGGGT	(SEQ ID NO:662);	GTACTGGGTC	(SEQ ID NO:663);
TACTGGGTCT	(SEQ ID NO:664);	ACTGGGTCTC	(SEQ ID NO:665);
CTGGGTCTCT	(SEQ ID NO:666);	TGGGTCTCTC	(SEQ ID NO:667);
GGGTCTCTCT	(SEQ ID NO:668);	GGTCTCTCTG	(SEQ ID NO:669);
GTCTCTCTGG	(SEQ ID NO:670);	TCTCTCTGGT	(SEQ ID NO:671);
CTCTCTGGTT	(SEQ ID NO:672);	TCTCTGGTTA	(SEQ ID NO:673);
CTCTGGTTAG	(SEQ ID NO:674);	TCTCTGGTTA	(SEQ ID NO:675);
CTGGTTAGAC	(SEQ ID NO:676);	TGGTTAGACC	(SEQ ID NO:677);
GGTTAGACCA	(SEQ ID NO:678);	GTTAGACCAG	(SEQ ID NO:679);
TTAGACCAGA	(SEQ ID NO:680);	TAGACCAGAT	(SEQ ID NO:681);
AGACCAGATC	(SEQ ID NO:682);	GACCAGATCT	(SEQ ID NO:683);
ACCAGATCTG	(SEQ ID NO:684);	CCAGATCTGA	(SEQ ID NO:685);
CAGATCTGAG	(SEQ ID NO:686);	AGATCTGAGC	(SEQ ID NO:687);
GATCTGAGCC	(SEQ ID NO:688);	ATCTGAGCCT	(SEQ ID NO:689);
TCTGAGCCTG	(SEQ ID NO:690);	CTGAGCCTGG	(SEQ ID NO:691);
TGAGCCTGGG	(SEQ ID NO:692);	GAGCCTGGGA	(SEQ ID NO:693);
AGCCTGGGAG	(SEQ ID NO:694);	GCCTGGGAGC	(SEQ ID NO:695);
CCTGGGAGCT	(SEQ ID NO:696);	CTGGGAGCTC	(SEQ ID NO:697);
TGGGAGCTCT	(SEQ ID NO:698);	GGGAGCTCTC	(SEQ ID NO:699);
GGAGCTCTCT	(SEQ ID NO:700);	GAGCTCTCTG	(SEQ ID NO:701);
AGCTCTCTGG	(SEQ ID NO:702);	GCTCTCTGGC	(SEQ ID NO:703);
CTCTCTGGCT	(SEQ ID NO:704);	TCTCTGGCTA	(SEQ ID NO:705);
CTCTGGCTAA	(SEQ ID NO:706);	TCTGGCTAAC	(SEQ ID NO:707);
CTGGCTAACT	(SEQ ID NO:708);	TGGCTAACTA	(SEQ ID NO:709);
GGCTAACTAG	(SEQ ID NO:710);	GCTAACTAGG	(SEQ ID NO:711);
CTAACTAGGG	(SEQ ID NO:712);	TAACTAGGGA	(SEQ ID NO:713);
AACTAGGGAA	(SEQ ID NO:714);	ACTAGGGAAC	(SEQ ID NO:715);
CTAGGGGAACC	(SEQ ID NO:716);	TAGGGGAACCC	(SEQ ID NO:717);
AGGGGAACCCA	(SEQ ID NO:718);	GGGAACCCAC	(SEQ ID NO:719);

- 210 -

GGAACCCACT	(SEQ ID NO:720);	GAACCCACTG	(SEQ ID NO:721);
AACCCACTGC	(SEQ ID NO:722);	ACCCACTGCT	(SEQ ID NO:723);
CCCACTGCTT	(SEQ ID NO:724);	CCACTGCTTA	(SEQ ID NO:725);
CACTGCTTAA	(SEQ ID NO:726);	ACTGCTTAAG	(SEQ ID NO:727);
CTGCTTAAGC	(SEQ ID NO:728);	TGCTTAAGCC	(SEQ ID NO:729);
GCTTAAGCCT	(SEQ ID NO:730);	CTTAAGCCTC	(SEQ ID NO:731);
TTAAGCCTCA	(SEQ ID NO:732);	TAAGCCTCAA	(SEQ ID NO:733);
AAGCCTCAAT	(SEQ ID NO:734);	AGCCTCAATA	(SEQ ID NO:735);
GCCTCAATAA	(SEQ ID NO:736);	CCTCAATAAA	(SEQ ID NO:737);
CTCAATAAAG	(SEQ ID NO:738);	TCAATAAAGC	(SEQ ID NO:739);
CAATAAAGCT	(SEQ ID NO:740);	AATAAAGCTT	(SEQ ID NO:741);
ATAAAGCTTG	(SEQ ID NO:742);	TAAAGCTTGC	(SEQ ID NO:743);
AAAGCTTGCC	(SEQ ID NO:744);	AAGCTTGCCCT	(SEQ ID NO:745);
AGCTTGCCCTT	(SEQ ID NO:746);	GCTTGCCCTTG	(SEQ ID NO:747);
CTTGCCCTGA	(SEQ ID NO:748);	TTGCCCTTGAG	(SEQ ID NO:749);
TGCCCTTGAGT	(SEQ ID NO:750);	GCCTTGAGTG	(SEQ ID NO:751);
CCTTGAGTGC	(SEQ ID NO:752);	CTTGAGTGCT	(SEQ ID NO:753);
TTGAGTGCTT	(SEQ ID NO:754);	TGAGTGCTTC	(SEQ ID NO:755);
GAGTGCTTCA	(SEQ ID NO:756);	AGTGCTTCAA	(SEQ ID NO:757);
GTGCTTCAAG	(SEQ ID NO:758);	TGCTTCAAGT	(SEQ ID NO:759);
GCTTCAAGTA	(SEQ ID NO:760);	CTTCAAGTAG	(SEQ ID NO:761);
TTCAAGTAGT	(SEQ ID NO:762);	TCAAGTAGTG	(SEQ ID NO:763);
CAAGTAGTGT	(SEQ ID NO:764);	AAGTAGTGTG	(SEQ ID NO:765);
AGTAGTGTGT	(SEQ ID NO:766);	GTAGTGTGTG	(SEQ ID NO:767);
TAGTGTGTGC	(SEQ ID NO:768);	AGTGTGTGCC	(SEQ ID NO:769);
GTGTGTGCCC	(SEQ ID NO:770);	TGTGTGCCCCG	(SEQ ID NO:771);
GTGTGCCCCGT	(SEQ ID NO:772);	TGTGCCCCGTC	(SEQ ID NO:773);
GTGCCCCGTCT	(SEQ ID NO:774);	TGCCCCGTCTG	(SEQ ID NO:775);
GCCCCGTCTGT	(SEQ ID NO:776);	CCCGTCTGTT	(SEQ ID NO:777);
CCGTCTGTTG	(SEQ ID NO:778);	CGTCTGTTGT	(SEQ ID NO:779);
GTCTGTTGTG	(SEQ ID NO:780);	TCTGTTGTGT	(SEQ ID NO:781);
CTGTTGTGTG	(SEQ ID NO:782);	TGTTGTGTGA	(SEQ ID NO:783);
GTTGTGTGAC	(SEQ ID NO:784);	TTGTGTGACT	(SEQ ID NO:785);
TGTGTGACTC	(SEQ ID NO:786);	GTGTGACTCT	(SEQ ID NO:787);
TGTGTGACTC	(SEQ ID NO:788);	GTGTGACTCT	(SEQ ID NO:789);
TGTGACTCTG	(SEQ ID NO:790);	GTGACTCTGG	(SEQ ID NO:791);
TGACTCTGGT	(SEQ ID NO:792);	GACTCTGGTA	(SEQ ID NO:793);
ACTCTGGTAA	(SEQ ID NO:794);	CTCTGGTAAC	(SEQ ID NO:795);



- 211 -

TCTGGTAACT (SEQ ID NO:796); CTGGTAACTA (SEQ ID NO:797);  
TGGTAACTAG (SEQ ID NO:798); GGTAAC TAGA (SEQ ID NO:799).

34. A method according to claim 24 wherein said HIV-1 isolate has the identifying characteristics of HIV-1 isolate ECACC V94101706 or HIV-1 isolate ECACC V941031169 or isolate C54 deposited at ECACC on 14 February, 1995 under Accession Number \_\_\_\_\_.

35. A method according to claim 20 wherein the HIV-1 isolate is reactive to antibodies to a glycoprotein of HIV-1, is capable of inducing an immune response of *gag*, *pol* and/or *env* and which is incapable of directing synthesis of a *nef* gene product or a full length *nef* gene product.

36. A method according to claim 35 wherein the glycoprotein is at least one of gp41-45, gp120 and/or gp160.

37. A method for vaccinating an individual against the development of AIDS or AIDS related diseases, said method comprising administering to said individual a non-pathogenic isolate of HIV-1 in an amount effective to infect target cells and to generate target cells carrying DNA derived from said non-pathogenic HIV-1.

38. A method according to claim 37 wherein said isolate is capable of stimulating an immune response in a human or primate against at least one glycoprotein on HIV-1 whilst not substantially reducing proliferative responses and cytokine production to a mitogen in said human or primate compared to a healthy, non-infected human or primate subject.

39. A method according to claim 38 wherein said isolate carries one or more mutations in its genome resulting in the inability to direct synthesis of at least one polypeptide or protein from a pathogenic strain of HIV-1.

40. A method according to claim 17 wherein said isolate carried one or more mutations in its genome resulting in said genome directing synthesis of a truncated form of a polypeptide or protein from a pathogenic strain of HIV-1.

41. A method according to claim 39 or 40 wherein said HIV-1 isolate carries a mutation in its *nef* gene and/or long terminal repeat (LTR) region or in a functionally equivalent location in the HIV-1 genome.

42. A method according to claim 40 wherein said HIV-1 isolate is reactive to antibodies to gp41-45, gp120 and/or gp160 of HIV-1 and carries a deletion of at least ten contiguous nucleotides in a region corresponding to the *nef* gene and/or LTR region of pathogenic HIV-1.

43. A method according to claim 42 wherein said HIV-1 isolate is capable of inducing an immune response to at least one of *gag*, *pol* and/or *env*.

44. A method according to claim 43 wherein said HIV-1 isolate carries a deletion in its genome of at least 10 nucleotides within the region from nucleotide 8787 to nucleotide 9709 using the nucleotide numbering of HIV-1 strain NL4-3.

45. A method according to claim 44 wherein said isolate carries a deletion in its genome of at least 10 nucleotides within a region selected from the list consisting of:

- |            |       |                |
|------------|-------|----------------|
| nucleotide | (i)   | 8830-8862;     |
|            | (ii)  | 9009-9035;     |
|            | (iii) | 9019-9029; and |
|            | (iv)  | 9033-9049.     |

46. A method according to claim 44 wherein said isolate carries a deletion in its genome of at least 10 nucleotides within a region selected from the list consisting of:

- |            |      |            |
|------------|------|------------|
| nucleotide | (v)  | 9281-9371; |
|            | (vi) | 9281-9362; |

- 213 -

- (vii) 9105-9224; and
- (viii) 9271-9370.

47. A method according to claim 44 wherein said isolate carries a deletion in its genome of at least 10 nucleotides within a region selected from the list consisting of:

- nucleotide (ix) 8882-8928;
- (x) 8850-9006;
- (xi) 8792-9041; and
- (xii) 9112-9204.

47. A method according to claim 44 wherein said isolate carries a deletion in its genome of at least 10 nucleotides within a region selected from the list consisting of:

- nucleotide (xiii) 9105-9224;
- (xiv) 9389-9395; and
- (xv) 9281-9366.

48. A method according to claim 44 wherein said deletion encompasses at least one of the following decanucleotides from the nef gene of HIV-1 strain NL43:

- |                            |                            |
|----------------------------|----------------------------|
| ATGGGTGGCA (SEQ ID NO:2);  | TGGGTGGCAA (SEQ ID NO:3);  |
| GGGTGGCAAG (SEQ ID NO:4);  | GGTGGCAAGT (SEQ ID NO:5);  |
| GTGGCAAGTG (SEQ ID NO:6);  | TGGCAAGTGG (SEQ ID NO:7);  |
| GGCAAGTGGT (SEQ ID NO:8);  | GCAAGTGGTC (SEQ ID NO:9);  |
| CAAGTGGTCA (SEQ ID NO:10); | AAGTGGTCAA (SEQ ID NO:11); |
| AGTGGTCAAA (SEQ ID NO:12); | GTGGTCAAAA (SEQ ID NO:13); |
| TGGTCAAAA (SEQ ID NO:14);  | GGTCAAAAAG (SEQ ID NO:15); |
| GTCAAAAAGT (SEQ ID NO:16); | TCAAAAAGTA (SEQ ID NO:17); |
| CAAAAAGTAG (SEQ ID NO:18); | AAAAAGTAGT (SEQ ID NO:19); |
| AAAAGTAGTG (SEQ ID NO:20); | AAAGTAGTGT (SEQ ID NO:21); |
| AAGTAGTGTG (SEQ ID NO:22); | AGTAGTGTGA (SEQ ID NO:23); |
| GTAGTGTGAT (SEQ ID NO:24); | TAGTGTGATT (SEQ ID NO:25); |
| AGTGTGATTG (SEQ ID NO:26); | GTGTGATTGG (SEQ ID NO:27); |
| TGTGATTGGA (SEQ ID NO:28); | GTGATTGGAT (SEQ ID NO:29); |
| TGATTGGATG (SEQ ID NO:30); | GATTGGATGG (SEQ ID NO:31); |
| ATTGGATGGC (SEQ ID NO:32); | TTGGATGGCC (SEQ ID NO:33); |
| TGGATGGCCT (SEQ ID NO:34); | GGATGGCCTG (SEQ ID NO:35); |

- 214 -

GATGGCCTGC (SEQ ID NO:36);	ATGGCCTGCT (SEQ ID NO:37);
TGGCCTGCTG (SEQ ID NO:38);	GGCCTGCTGT (SEQ ID NO:39);
GCCTGCTGTA (SEQ ID NO:40);	CCTGCTGTAA (SEQ ID NO:41);
CTGCTGTAAG (SEQ ID NO:42);	TGCTGTAAGG (SEQ ID NO:43);
GCTGTAAGGG (SEQ ID NO:44);	CTGTAAGGGA (SEQ ID NO:45);
TGTAAGGGAA (SEQ ID NO:46);	GTAAGGGAAA (SEQ ID NO:47);
TAAGGGAAAG (SEQ ID NO:48);	AAGGGAAAGA (SEQ ID NO:49);
AGGGAAAGAA (SEQ ID NO:50);	GGGAAAGAAT (SEQ ID NO:51);
GGAAAGAATG (SEQ ID NO:52);	GAAAGAATGA (SEQ ID NO:53);
AAAGAATGAG (SEQ ID NO:54);	AAGAATGAGA (SEQ ID NO:55);
AGAATGAGAC (SEQ ID NO:56);	GAATGAGACG (SEQ ID NO:57);
AATGAGACGA (SEQ ID NO:58);	ATGAGACGAG (SEQ ID NO:59);
TGAGACGAGC (SEQ ID NO:60);	GAGACGAGCT (SEQ ID NO:61);
AGACGAGCTG (SEQ ID NO:62);	GACGAGCTGA (SEQ ID NO:63);
ACGAGCTGAG (SEQ ID NO:64);	CGAGCTGAGC (SEQ ID NO:65);
GAGCTGAGCC (SEQ ID NO:66);	AGCTGAGCCA (SEQ ID NO:67);
GCTGAGCCAG (SEQ ID NO:68);	CTGAGCCAGC (SEQ ID NO:69);
TGAGCCAGCA (SEQ ID NO:70);	GAGCCAGCAG (SEQ ID NO:71);
AGCCAGCAGC (SEQ ID NO:72);	GCCAGCAGCA (SEQ ID NO:73);
CCAGCAGCAG (SEQ ID NO:74);	CAGCAGCAGA (SEQ ID NO:75);
AGCAGCAGAT (SEQ ID NO:76);	GCAGCAGATG (SEQ ID NO:77);
CAGCAGATGG (SEQ ID NO:78);	AGCAGATGGG (SEQ ID NO:79);
GCAGATGGGG (SEQ ID NO:80);	CAGATGGGGT (SEQ ID NO:81);
AGATGGGGTG (SEQ ID NO:82);	GATGGGGTGG (SEQ ID NO:83);
ATGGGGTGGG (SEQ ID NO:84);	TGGGGTGGGA (SEQ ID NO:85);
GGGGTGGGAG (SEQ ID NO:86);	GGGTGGGAGC (SEQ ID NO:87);
GGTGGGAGCA (SEQ ID NO:88);	GTGGGAGCAG (SEQ ID NO:89);
TGGGAGCAGT (SEQ ID NO:90);	GGGAGCAGTA (SEQ ID NO:91);
GGAGCAGTAT (SEQ ID NO:92);	GAGCAGTATC (SEQ ID NO:93);
AGCAGTATCT (SEQ ID NO:94);	GCACTATCTC (SEQ ID NO:95);
CAGTATCTCG (SEQ ID NO:96);	AGTATCTCGA (SEQ ID NO:97);
GTATCTCGAG (SEQ ID NO:98);	TATCTCGAGA (SEQ ID NO:99);
ATCTCGAGAC (SEQ ID NO:100);	TCTCGAGACC (SEQ ID NO:101);
CTCGAGACCT (SEQ ID NO:102);	TCGAGACCTA (SEQ ID NO:103);
CGAGACCTAG (SEQ ID NO:104);	GAGACCTAGA (SEQ ID NO:105);
AGACCTAGAA (SEQ ID NO:106);	GACCTAGAAA (SEQ ID NO:107);
ACCTAGAAAA (SEQ ID NO:108);	CCTAGAAAAA (SEQ ID NO:109);
CTAGAAAAAC (SEQ ID NO:110);	TAGAAAAACA (SEQ ID NO:111);

- 215 -

AGAAAAACAT (SEQ ID NO:112);	GAAAAACATG (SEQ ID NO:113);
AAAAACATGG (SEQ ID NO:114);	AAAACATGGA (SEQ ID NO:115);
AAACATGGAG (SEQ ID NO:116);	AACATGGAGC (SEQ ID NO:117);
ACATGGAGCA (SEQ ID NO:118);	CATGGAGCAA (SEQ ID NO:119);
ATGGAGCAAT (SEQ ID NO:120);	TGGAGCAATC (SEQ ID NO:121);
GGAGCAATCA (SEQ ID NO:122);	GAGCAATCAC (SEQ ID NO:123);
AGCAATCACA (SEQ ID NO:124);	GCAATCACAA (SEQ ID NO:125);
CAATCACAAG (SEQ ID NO:126);	AATCACAAGT (SEQ ID NO:127);
ATCACAAGTA (SEQ ID NO:128);	TCACAAGTAG (SEQ ID NO:129);
CACAAGTAGC (SEQ ID NO:130);	ACAAGTAGCA (SEQ ID NO:131);
CAAGTAGCAA (SEQ ID NO:132);	AAGTAGCAAT (SEQ ID NO:133);
AGTAGCAATA (SEQ ID NO:134);	GTAGCAATAC (SEQ ID NO:135);
TAGCAATACA (SEQ ID NO:136);	AGCAATACAG (SEQ ID NO:137);
GCAATACAGC (SEQ ID NO:138);	CAATACAGCA (SEQ ID NO:139);
AATACAGCAG (SEQ ID NO:140);	ATACAGCAGC (SEQ ID NO:141);
TACAGCAGCT (SEQ ID NO:142);	ACAGCAGCTA (SEQ ID NO:143);
CAGCAGCTAA (SEQ ID NO:144);	AGCAGCTAAC (SEQ ID NO:145);
GCAGCTAACA (SEQ ID NO:146);	CAGCTAACAA (SEQ ID NO:147);
AGCTAACAAT (SEQ ID NO:148);	GCTAACAATG (SEQ ID NO:149);
CTAACAATGC (SEQ ID NO:150);	TAACAATGCT (SEQ ID NO:151);
AACAATGCTG (SEQ ID NO:152);	ACAATGCTGC (SEQ ID NO:153);
CAATGCTGCT (SEQ ID NO:154);	AATGCTGCTT (SEQ ID NO:155);
ATGCTGCTTG (SEQ ID NO:156);	TGCTGCTTGT (SEQ ID NO:157);
GCTGCTTGTG (SEQ ID NO:158);	CTGCTTGTGC (SEQ ID NO:159);
TGCTTGTGCC (SEQ ID NO:160);	GCTTGTGCCT (SEQ ID NO:161);
CTTGTGCCTG (SEQ ID NO:162);	TTGTGCCTGG (SEQ ID NO:163);
TGTGCCTGGC (SEQ ID NO:164);	GTGCCTGGCT (SEQ ID NO:165);
TGCCTGGCTA (SEQ ID NO:166);	GCCTGGCTAG (SEQ ID NO:167);
CCTGGCTAGA (SEQ ID NO:168);	CTGGCTAGAA (SEQ ID NO:169);
TGGCTAGAAG (SEQ ID NO:170);	GGCTAGAAGC (SEQ ID NO:171);
GCTAGAAGCA (SEQ ID NO:172);	CTAGAAGCAC (SEQ ID NO:173);
TAGAAGCACA (SEQ ID NO:174);	AGAAGCACAA (SEQ ID NO:175);
GAAGCACAAG (SEQ ID NO:176);	AAGCACAAGA (SEQ ID NO:177);
AGCACAAGAG (SEQ ID NO:178);	GCACAAGAGG (SEQ ID NO:179);
CACAAGAGGA (SEQ ID NO:180);	ACAAGAGGAG (SEQ ID NO:181);
CAAGAGGAGG (SEQ ID NO:182);	AAGAGGAGGA (SEQ ID NO:183);
AGAGGAGGAA (SEQ ID NO:184);	GAGGAGGAAG (SEQ ID NO:185);
AGGAGGAAGA (SEQ ID NO:186);	GGAGGAAGAG (SEQ ID NO:187);

- 216 -

GAGGAAGAGG (SEQ ID NO:188);	AGGAAGAGGT (SEQ ID NO:189);
GGAAGAGGTG (SEQ ID NO:190);	GAAGAGGTGG (SEQ ID NO:191);
AAGAGGTGGG (SEQ ID NO:192);	AGAGGTGGGT (SEQ ID NO:193);
GAGGTGGGTT (SEQ ID NO:194);	AGGTGGGTTT (SEQ ID NO:195);
GGTGGGTTTT (SEQ ID NO:196);	GTGGGTTTTTC (SEQ ID NO:197);
TGGGTTTTTCC (SEQ ID NO:198);	GGGTTTTTCCA (SEQ ID NO:199);
GGTTTTTCCAG (SEQ ID NO:200);	GTTTTTCCAGT (SEQ ID NO:201);
TTTTTCCAGTC (SEQ ID NO:202);	TTTCCAGTCA (SEQ ID NO:203);
TTCCAGTCAC (SEQ ID NO:204);	TCCAGTCACA (SEQ ID NO:205);
CCAGTCACAC (SEQ ID NO:206);	CAGTCACACC (SEQ ID NO:207);
AGTCACACCT (SEQ ID NO:208);	GTCACACCTC (SEQ ID NO:209);
TCACACCTCA (SEQ ID NO:210);	CACACCTCAG (SEQ ID NO:211);
ACACCTCAGG (SEQ ID NO:212);	CACCTCAGGT (SEQ ID NO:213);
ACCTCAGGTA (SEQ ID NO:214);	CCTCAGGTAC (SEQ ID NO:215);
CTCAGGTACC (SEQ ID NO:216);	TCAGGTACCT (SEQ ID NO:217);
CAGGTACCTT (SEQ ID NO:218);	AGGTACCTTT (SEQ ID NO:219);
GGTACCTTTA (SEQ ID NO:220);	GTACCTTTAA (SEQ ID NO:221);
TACCTTTAAG (SEQ ID NO:222);	ACCTTTAAGA (SEQ ID NO:223);
CCTTTAAGAC (SEQ ID NO:224);	CTTTAAGACC (SEQ ID NO:225);
TTTAAGACCA (SEQ ID NO:226);	TTAAGACCAA (SEQ ID NO:227);
TAAGACCAAT (SEQ ID NO:228);	AAGACCAATG (SEQ ID NO:229);
AGACCAATGA (SEQ ID NO:230);	GACCAATGAC (SEQ ID NO:231);
ACCAATGACT (SEQ ID NO:232);	CCAATGACTT (SEQ ID NO:233);
CAATGACTTA (SEQ ID NO:234);	AATGACTTAC (SEQ ID NO:235);
ATGACTTACA (SEQ ID NO:236);	TGACTTACAA (SEQ ID NO:237);
GACTTACAAG (SEQ ID NO:238);	ACTTACAAGG (SEQ ID NO:239);
CTTACAAGGC (SEQ ID NO:240);	TTACAAGGCA (SEQ ID NO:241);
TACAAGGCAG (SEQ ID NO:242);	ACAAGGCAGC (SEQ ID NO:243);
CAAGGCAGCT (SEQ ID NO:244);	AAGGCAGCTG (SEQ ID NO:245);
AGGCAGCTGT (SEQ ID NO:246);	GGCAGCTGTA (SEQ ID NO:247);
GCAGCTGTAG (SEQ ID NO:248);	CAGCTGTAGA (SEQ ID NO:249);
AGCTGTAGAT (SEQ ID NO:250);	GCTGTAGATC (SEQ ID NO:251);
CTGTAGATCT (SEQ ID NO:252);	TGTAGATCTT (SEQ ID NO:253);
GTAGATCTTA (SEQ ID NO:254);	TAGATCTTAG (SEQ ID NO:255);
AGATCTTAGC (SEQ ID NO:256);	GATCTTAGCC (SEQ ID NO:257);
ATCTTAGCCA (SEQ ID NO:258);	TCTTAGCCAC (SEQ ID NO:259);
CTTAGCCACT (SEQ ID NO:260);	TTAGCCACTT (SEQ ID NO:261);
TAGCCACTTT (SEQ ID NO:262);	AGCCACTTTT (SEQ ID NO:263);

- 217 -

GCCACTTTTTT (SEQ ID NO:264);	CCACTTTTTTA (SEQ ID NO:265);
CACTTTTTTAA (SEQ ID NO:266);	ACTTTTTTAAA (SEQ ID NO:267);
CTTTTTTAAAA (SEQ ID NO:268);	TTTTTTAAAAG (SEQ ID NO:269);
TTTTTAAAGA (SEQ ID NO:270);	TTTAAAGAA (SEQ ID NO:271);
TTAAAGAAA (SEQ ID NO:272);	TAAAGAAAA (SEQ ID NO:273);
AAAAGAAAAG (SEQ ID NO:274);	AAAGAAAAGG (SEQ ID NO:275);
AAGAAAAGGG (SEQ ID NO:276);	AGAAAAGGGG (SEQ ID NO:277);
GAAAAGGGGG (SEQ ID NO:278);	AAAAGGGGGG (SEQ ID NO:279);
AAAGGGGGGA (SEQ ID NO:280);	AAGGGGGGAC (SEQ ID NO:281);
AGGGGGGACT (SEQ ID NO:282);	GGGGGGACTG (SEQ ID NO:283);
GGGGGACTGG (SEQ ID NO:284);	GGGGACTGGA (SEQ ID NO:285);
GGGACTGGAA (SEQ ID NO:286);	GGACTGGAAG (SEQ ID NO:287);
GACTGGAAGG (SEQ ID NO:288);	ACTGGAAGGG (SEQ ID NO:289);
CTGGAAGGGC (SEQ ID NO:290);	TGGAAGGGCT (SEQ ID NO:291);
GGAAGGGCTA (SEQ ID NO:292);	GAAGGGCTAA (SEQ ID NO:293);
AAGGGCTAAT (SEQ ID NO:294);	AGGGCTAATT (SEQ ID NO:295);
GGGCTAATTC (SEQ ID NO:296);	GGCTAATTCA (SEQ ID NO:297);
GCTAATTCAC (SEQ ID NO:298);	CTAATTCACT (SEQ ID NO:299);
TAATTCACTC (SEQ ID NO:300);	AATTCACTCC (SEQ ID NO:301);
ATTCACTCCC (SEQ ID NO:302);	TTCACTCCCA (SEQ ID NO:303);
TCACTCCCAA (SEQ ID NO:304);	CACTCCCAAA (SEQ ID NO:305);
ACTCCCAAAG (SEQ ID NO:306);	CTCCCAAAGA (SEQ ID NO:307);
TCCCAAAGAA (SEQ ID NO:308);	CCCAAAGAAG (SEQ ID NO:309);
CCAAAGAAGA (SEQ ID NO:310);	CAAAGAAGAC (SEQ ID NO:311);
AAAGAAGACA (SEQ ID NO:312);	AAGAAGACAA (SEQ ID NO:313);
AGAAGACAAG (SEQ ID NO:314);	GAAGACAAGA (SEQ ID NO:315);
AAGACAAGAT (SEQ ID NO:316);	AGACAAGATA (SEQ ID NO:317);
GACAAGATAT (SEQ ID NO:318);	ACAAGATATC (SEQ ID NO:319);
CAAGATATCC (SEQ ID NO:320);	AAGATATCCT (SEQ ID NO:321);
AGATATCCTT (SEQ ID NO:322);	GATATCCTTG (SEQ ID NO:323);
ATATCCTTGA (SEQ ID NO:324);	TATCCTTGAT (SEQ ID NO:325);
ATCCTTGATC (SEQ ID NO:326);	TCCTTGATCT (SEQ ID NO:327);
CCTTGATCTG (SEQ ID NO:328);	CTTGATCTGT (SEQ ID NO:329);
TTGATCTGTG (SEQ ID NO:330);	TGATCTGTGG (SEQ ID NO:331);
GATCTGTGGA (SEQ ID NO:332);	ATCTGTGGAT (SEQ ID NO:333);
TCTGTGGATC (SEQ ID NO:334);	CTGTGGATCT (SEQ ID NO:335);
TGTGGATCTA (SEQ ID NO:336);	GTGGATCTAC (SEQ ID NO:337);
TGGATCTACC (SEQ ID NO:338);	GGATCTACCA (SEQ ID NO:339);

- 218 -

GATCTACCAC (SEQ ID NO:340);	ATCTACCACA (SEQ ID NO:341);
TCTACCACAC (SEQ ID NO:342);	CTACCACACA (SEQ ID NO:343);
TACCACACAC (SEQ ID NO:344);	ACCACACACA (SEQ ID NO:345);
CCACACACAA (SEQ ID NO:346);	CACACACAAG (SEQ ID NO:347);
ACACACAAGG (SEQ ID NO:348);	CACACAAGGC (SEQ ID NO:349);
ACACAAGGCT (SEQ ID NO:350);	CACAAGGCTA (SEQ ID NO:351);
ACAAGGCTAC (SEQ ID NO:352);	CAAGGCTACT (SEQ ID NO:353);
AAGGCTACTT (SEQ ID NO:354);	AGGCTACTTC (SEQ ID NO:355);
GGCTACTTCC (SEQ ID NO:356);	GCTACTTCCC (SEQ ID NO:357);
CTACTTCCCT (SEQ ID NO:358);	TACTTCCCTG (SEQ ID NO:359);
ACTTCCCTGA (SEQ ID NO:360);	CTTCCCTGAT (SEQ ID NO:361);
TTCCCTGATT (SEQ ID NO:362);	TCCCTGATTG (SEQ ID NO:363);
CCCTGATTGG (SEQ ID NO:364);	CCTGATTGGC (SEQ ID NO:365);
CTGATTGGCA (SEQ ID NO:366);	TGATTGGCAG (SEQ ID NO:367);
GATTGGCAGA (SEQ ID NO:368);	ATTGGCAGAA (SEQ ID NO:369);
TTGGCAGAAC (SEQ ID NO:370);	TGGCAGAACT (SEQ ID NO:371);
GGCAGAACTA (SEQ ID NO:372);	GCAGAACTAC (SEQ ID NO:373);
CAGAACTACA (SEQ ID NO:374);	AGAACTACAC (SEQ ID NO:375);
GAACTACACA (SEQ ID NO:376);	AACTACACAC (SEQ ID NO:377);
ACTACACACC (SEQ ID NO:378);	CTACACACCA (SEQ ID NO:379);
TACACACCAG (SEQ ID NO:380);	ACACACCAGG (SEQ ID NO:381);
CACACCAGGG (SEQ ID NO:382);	ACACCAGGGC (SEQ ID NO:383);
CACCAGGGCC (SEQ ID NO:384);	ACCAGGGCCA (SEQ ID NO:385);
CCAGGGCCAG (SEQ ID NO:386);	CAGGGCCAGG (SEQ ID NO:387);
AGGGCCAGGG (SEQ ID NO:388);	GGGCCAGGGG (SEQ ID NO:389);
GGCCAGGGGT (SEQ ID NO:390);	GCCAGGGGTC (SEQ ID NO:391);
CCAGGGGTCA (SEQ ID NO:392);	CAGGGGTCAG (SEQ ID NO:393);
AGGGGTCA (SEQ ID NO:394);	GGGGTCAGAT (SEQ ID NO:395);
GGGTCA (SEQ ID NO:396);	GGTCAGATAT (SEQ ID NO:397);
GTCAGATATC (SEQ ID NO:398);	TCAGATATCC (SEQ ID NO:399);
CAGATATCCA (SEQ ID NO:400);	AGATATCCAC (SEQ ID NO:401);
GATATCCACT (SEQ ID NO:402);	ATATCCACTG (SEQ ID NO:403);
TATCCACTGA (SEQ ID NO:404);	ATCCACTGAC (SEQ ID NO:405);
TCCACTGACC (SEQ ID NO:406);	CCACTGACCT (SEQ ID NO:407);
CACTGACCTT (SEQ ID NO:408);	ACTGACCTTT (SEQ ID NO:409);
CTGACCTTTG (SEQ ID NO:410);	TGACCTTTGG (SEQ ID NO:411);
GACCTTTGGA (SEQ ID NO:412);	ACCTTTGGAT (SEQ ID NO:413);
CCTTTGGATG (SEQ ID NO:414);	CTTTGGATGG (SEQ ID NO:415);



- 219 -

TTTGATGGT (SEQ ID NO:416);	TTGGATGGTG (SEQ ID NO:417);
TGGATGGTGC (SEQ ID NO:418);	GGATGGTGCT (SEQ ID NO:419);
GATGGTGCTA (SEQ ID NO:420);	ATGGTGCTAC (SEQ ID NO:421);
TGGTGCTACA (SEQ ID NO:422);	GGTGCTACAA (SEQ ID NO:423);
GTGCTACAAG (SEQ ID NO:424);	TGCTACAAGC (SEQ ID NO:425);
GCTACAAGCT (SEQ ID NO:426);	CTACAAGCTA (SEQ ID NO:427);
TACAAGCTAG (SEQ ID NO:428);	ACAAGCTAGT (SEQ ID NO:429);
CAAGCTAGTA (SEQ ID NO:430);	AAGCTAGTAC (SEQ ID NO:431);
AGCTAGTACC (SEQ ID NO:432);	GCTAGTACCA (SEQ ID NO:433);
CTAGTACCAG (SEQ ID NO:434);	TAGTACCAGT (SEQ ID NO:435);
AGTACCAGTT (SEQ ID NO:436);	GTACCAGTTG (SEQ ID NO:437);
TACCAGTTGA (SEQ ID NO:438);	ACCAGTTGAG (SEQ ID NO:439);
CCAGTTGAGC (SEQ ID NO:440);	CAGTTGAGCC (SEQ ID NO:441);
AGTTGAGCCA (SEQ ID NO:442);	GTTGAGCCAG (SEQ ID NO:443);
TTGAGCCAGA (SEQ ID NO:444);	TGAGCCAGAT (SEQ ID NO:445);
GAGCCAGATA (SEQ ID NO:446);	AGCCAGATAA (SEQ ID NO:447);
GCCAGATAAG (SEQ ID NO:448);	CCAGATAAGG (SEQ ID NO:449);
CAGATAAGGT (SEQ ID NO:450);	AGATAAGGTA (SEQ ID NO:451);
GATAAGGTAG (SEQ ID NO:452);	ATAAGGTAGA (SEQ ID NO:453);
TAAGGTAGAA (SEQ ID NO:454);	AAGGTAGAAG (SEQ ID NO:455);
AGGTAGAAGA (SEQ ID NO:456);	GGTAGAAGAG (SEQ ID NO:457);
GTAGAAGAGG (SEQ ID NO:458);	TAGAAGAGGC (SEQ ID NO:459);
AGAAGAGGCC (SEQ ID NO:460);	GAAGAGGCCA (SEQ ID NO:461);
AAGAGGCCAA (SEQ ID NO:462);	AGAGGCCAAT (SEQ ID NO:463);
GAGGCCAATA (SEQ ID NO:464);	AGGCCAATAA (SEQ ID NO:465);
GGCCAATAAA (SEQ ID NO:466);	GCCAATAAAG (SEQ ID NO:467);
CCAATAAAGG (SEQ ID NO:468);	CAATAAAGGA (SEQ ID NO:469);
AATAAAGGAG (SEQ ID NO:470);	ATAAAGGAGA (SEQ ID NO:471);
TAAAGGAGAG (SEQ ID NO:472);	AAAGGAGAGA (SEQ ID NO:473);
AAGGAGAGAA (SEQ ID NO:474);	AGGAGAGAAC (SEQ ID NO:475);
GGAGAGAACA (SEQ ID NO:476);	GAGAGAACAC (SEQ ID NO:477);
AGAGAACACC (SEQ ID NO:478);	GAGAACACCA (SEQ ID NO:479);
AGAACACCAG (SEQ ID NO:480);	GAACACCAGC (SEQ ID NO:481);
AACACCAGCT (SEQ ID NO:482);	ACACCAGCTT (SEQ ID NO:483);
CACCAGCTTG (SEQ ID NO:484);	ACCAGCTTGT (SEQ ID NO:485);
CCAGCTTGTT (SEQ ID NO:486);	CAGCTTGTTA (SEQ ID NO:487);
AGCTTGTTAC (SEQ ID NO:488);	GCTTGTTACA (SEQ ID NO:489);
CTTGTTACAC (SEQ ID NO:490);	TTGTTACACC (SEQ ID NO:491);

TGTTACACCC (SEQ ID NO:492);	GTTACACCCT (SEQ ID NO:493);
TTACACCCTG (SEQ ID NO:494);	TACACCCTGT (SEQ ID NO:495);
ACACCCTGTG (SEQ ID NO:496);	CACCCTGTGA (SEQ ID NO:497);
ACCCTGTGAG (SEQ ID NO:498);	CCCTGTGAGC (SEQ ID NO:499);
CCTGTGAGCC (SEQ ID NO:500);	CTGTGAGCCT (SEQ ID NO:501);
TGTGAGCCTG (SEQ ID NO:502);	GTGAGCCTGC (SEQ ID NO:503);
TGAGCCTGCA (SEQ ID NO:504);	GAGCCTGCAT (SEQ ID NO:505);
AGCCTGCATG (SEQ ID NO:506);	GCCTGCATGG (SEQ ID NO:507);
CCTGCATGGA (SEQ ID NO:508);	CTGCATGGAA (SEQ ID NO:509);
TGCATGGAAT (SEQ ID NO:510);	GCATGGAATG (SEQ ID NO:511);
CATGGAATGG (SEQ ID NO:512);	ATGGAATGGA (SEQ ID NO:513);
TGGAATGGAT (SEQ ID NO:514);	GGAATGGATG (SEQ ID NO:515);
GAATGGATGA (SEQ ID NO:516);	AATGGATGAC (SEQ ID NO:517);
ATGGATGACC (SEQ ID NO:518);	TGGATGACCC (SEQ ID NO:519);
GGATGACCCT (SEQ ID NO:520);	GATGACCCTG (SEQ ID NO:521);
ATGACCCTGA (SEQ ID NO:522);	TGACCCTGAG (SEQ ID NO:523);
GACCCTGAGA (SEQ ID NO:524);	ACCCTGAGAG (SEQ ID NO:525);
CCCTGAGAGA (SEQ ID NO:526);	CCTGAGAGAG (SEQ ID NO:527);
CTGAGAGAGA (SEQ ID NO:528);	TGAGAGAGAA (SEQ ID NO:529);
GAGAGAGAAG (SEQ ID NO:530);	AGAGAGAAGT (SEQ ID NO:531);
GAGAGAAGTG (SEQ ID NO:532);	AGAGAAGTGT (SEQ ID NO:533);
GAGAAGTGTT (SEQ ID NO:534);	AGAAGTGTTA (SEQ ID NO:535);
GAAGTGTTAG (SEQ ID NO:536);	AAGTGTTAGA (SEQ ID NO:537);
AGTGTTAGAG (SEQ ID NO:538);	GTGTTAGAGT (SEQ ID NO:539);
TGTTAGAGTG (SEQ ID NO:540);	GTTAGAGTGG (SEQ ID NO:541);
TTAGAGTGGG (SEQ ID NO:542);	TAGAGTGGAG (SEQ ID NO:543);
AGAGTGGAGG (SEQ ID NO:544);	GAGTGGAGGT (SEQ ID NO:545);
AGTGGAGGTT (SEQ ID NO:546);	GTGGAGGTTT (SEQ ID NO:547);
TGGAGGTTTG (SEQ ID NO:548);	GGAGGTTTGA (SEQ ID NO:549);
GAGGTTTGAC (SEQ ID NO:550);	AGGTTTGACA (SEQ ID NO:551);
GGTTTGACAG (SEQ ID NO:552);	GTTTGACAGC (SEQ ID NO:553);
TTTGACAGCC (SEQ ID NO:554);	TTGACAGCCG (SEQ ID NO:555);
TGACAGCCGC (SEQ ID NO:556);	GACAGCCGCC (SEQ ID NO:557);
ACAGCCGCCT (SEQ ID NO:558);	CAGCCGCCTA (SEQ ID NO:559);
AGCCGCCTAG (SEQ ID NO:560);	GCCGCCTAGC (SEQ ID NO:561);
CCGCCTAGCA (SEQ ID NO:562);	CGCCTAGCAT (SEQ ID NO:563);
GCCTAGCATT (SEQ ID NO:564);	CCTAGCATTT (SEQ ID NO:565);
CTAGCATTTC (SEQ ID NO:566);	TAGCATTTCA (SEQ ID NO:567);

- 221 -

AGCATTTTCAT (SEQ ID NO:568);	GCATTTTCATC (SEQ ID NO:569);
CATTTTCATCA (SEQ ID NO:570);	ATTTTCATCAC (SEQ ID NO:571);
TTTCATCACG (SEQ ID NO:572);	TTCATCACGT (SEQ ID NO:573);
TCATCACGTG (SEQ ID NO:574);	CATCACGTGG (SEQ ID NO:575);
ATCACGTGGC (SEQ ID NO:576);	TCACGTGGCC (SEQ ID NO:577);
CACGTGGCCC (SEQ ID NO:578);	ACGTGGCCCCG (SEQ ID NO:579);
CGTGGCCCCG (SEQ ID NO:580);	GTGGCCCCGAG (SEQ ID NO:581);
TGGCCCCGAGA (SEQ ID NO:582);	GGCCCCGAGAG (SEQ ID NO:583);
GCCCCGAGAGC (SEQ ID NO:584);	CCCGAGAGCT (SEQ ID NO:585);
CCGAGAGCTG (SEQ ID NO:586);	CGAGAGCTGC (SEQ ID NO:587);
GAGAGCTGCA (SEQ ID NO:588);	AGAGCTGCAT (SEQ ID NO:589);
GAGCTGCATC (SEQ ID NO:590);	AGCTGCATCC (SEQ ID NO:591);
GCTGCATCCG (SEQ ID NO:592);	CTGCATCCGG (SEQ ID NO:593);
TGCATCCGGA (SEQ ID NO:594);	GCATCCGGAG (SEQ ID NO:595);
CATCCGGAGT (SEQ ID NO:596);	ATCCGGAGTA (SEQ ID NO:597);
TCCGGAGTAC (SEQ ID NO:598);	CCGGAGTACT (SEQ ID NO:599);
CGGAGTACTT (SEQ ID NO:600);	GGAGTACTTC (SEQ ID NO:601);
GAGTACTTCA (SEQ ID NO:602);	AGTACTTCAA (SEQ ID NO:603);
GTACTTCAAG (SEQ ID NO:604);	TACTTCAAGA (SEQ ID NO:605);
ACTTCAAGAA (SEQ ID NO:606);	CTTCAAGAAC (SEQ ID NO:607);
TTCAAGAACT (SEQ ID NO:608);	TCAAGAACTG (SEQ ID NO:609);
CAAGAACTGC (SEQ ID NO:610);	AAGAACTGCT (SEQ ID NO:611);
AGAAGCTGCTG (SEQ ID NO:612);	GAACTGCTGA (SEQ ID NO:613);

49. A method according to claim 44 wherein said deletion encompasses at least one of the following decanucleotides from the LTR region of HIV-1 strain NL4-3:

GCTTTTTGCC (SEQ ID NO:652);	CTTTTTGCCT (SEQ ID NO:653);
TTTTTGCCTG (SEQ ID NO:654);	TTTTGCCTGT (SEQ ID NO:655);
TTTGCCTGTA (SEQ ID NO:656);	TTGCCTGTAC (SEQ ID NO:657);
TGCCTGTACT (SEQ ID NO:658);	GCCTGTACTG (SEQ ID NO:659);
CCTGTACTGG (SEQ ID NO:660);	CTGTACTGGG (SEQ ID NO:661);
TGTACTGGGT (SEQ ID NO:662);	GTACTGGGTC (SEQ ID NO:663);
TACTGGGTCT (SEQ ID NO:664);	ACTGGGTCTC (SEQ ID NO:665);
CTGGGTCTCT (SEQ ID NO:666);	TGGGTCTCTC (SEQ ID NO:667);
GGGTCTCTCT (SEQ ID NO:668);	GGTCTCTCTG (SEQ ID NO:669);
GTCTCTCTGG (SEQ ID NO:670);	TCTCTCTGGT (SEQ ID NO:671);
CTCTCTGGTT (SEQ ID NO:672);	TCTCTGGTTA (SEQ ID NO:673);

- 222 -

CTCTGGTTAG	(SEQ ID NO:674);	TCTCTGGTTA	(SEQ ID NO:675);
CTGGTTAGAC	(SEQ ID NO:676);	TGGTTAGACC	(SEQ ID NO:677);
GGTTAGACCA	(SEQ ID NO:678);	GTTAGACCAG	(SEQ ID NO:679);
TTAGACCAGA	(SEQ ID NO:680);	TAGACCAGAT	(SEQ ID NO:681);
AGACCAGATC	(SEQ ID NO:682);	GACCAGATCT	(SEQ ID NO:683);
ACCAGATCTG	(SEQ ID NO:684);	CCAGATCTGA	(SEQ ID NO:685);
CAGATCTGAG	(SEQ ID NO:686);	AGATCTGAGC	(SEQ ID NO:687);
GATCTGAGCC	(SEQ ID NO:688);	ATCTGAGCCT	(SEQ ID NO:689);
TCTGAGCCTG	(SEQ ID NO:690);	CTGAGCCTGG	(SEQ ID NO:691);
TGAGCCTGGG	(SEQ ID NO:692);	GAGCCTGGGA	(SEQ ID NO:693);
AGCCTGGGAG	(SEQ ID NO:694);	GCCTGGGAGC	(SEQ ID NO:695);
CCTGGGAGCT	(SEQ ID NO:696);	CTGGGAGCTC	(SEQ ID NO:697);
TGGGAGCTCT	(SEQ ID NO:698);	GGGAGCTCTC	(SEQ ID NO:699);
GGAGCTCTCT	(SEQ ID NO:700);	GAGCTCTCTG	(SEQ ID NO:701);
AGCTCTCTGG	(SEQ ID NO:702);	GCTCTCTGGC	(SEQ ID NO:703);
CTCTCTGGCT	(SEQ ID NO:704);	TCTCTGGCTA	(SEQ ID NO:705);
CTCTGGCTAA	(SEQ ID NO:706);	TCTGGCTAAC	(SEQ ID NO:707);
CTGGCTAACT	(SEQ ID NO:708);	TGGCTAACTA	(SEQ ID NO:709);
GGCTAACTAG	(SEQ ID NO:710);	GCTAACTAGG	(SEQ ID NO:711);
CTAACTAGGG	(SEQ ID NO:712);	TAACTAGGGA	(SEQ ID NO:713);
AACTAGGGAA	(SEQ ID NO:714);	ACTAGGGAAC	(SEQ ID NO:715);
CTAGGGAAACC	(SEQ ID NO:716);	TAGGGAAACCC	(SEQ ID NO:717);
AGGGAACCCA	(SEQ ID NO:718);	GGGAACCCAC	(SEQ ID NO:719);
GGAACCCACT	(SEQ ID NO:720);	GAACCCACTG	(SEQ ID NO:721);
AACCCACTGC	(SEQ ID NO:722);	ACCCACTGCT	(SEQ ID NO:723);
CCCACTGCTT	(SEQ ID NO:724);	CCACTGCTTA	(SEQ ID NO:725);
CACTGCTTAA	(SEQ ID NO:726);	ACTGCTTAAG	(SEQ ID NO:727);
CTGCTTAAGC	(SEQ ID NO:728);	TGCTTAAGCC	(SEQ ID NO:729);
GCTTAAGCCT	(SEQ ID NO:730);	CTTAAGCCTC	(SEQ ID NO:731);
TTAAGCCTCA	(SEQ ID NO:732);	TAAGCCTCAA	(SEQ ID NO:733);
AAGCCTCAAT	(SEQ ID NO:734);	AGCCTCAATA	(SEQ ID NO:735);
GCCTCAATAA	(SEQ ID NO:736);	CCTCAATAAA	(SEQ ID NO:737);
CTCAATAAAG	(SEQ ID NO:738);	TCAATAAAGC	(SEQ ID NO:739);
CAATAAAGCT	(SEQ ID NO:740);	AATAAAGCTT	(SEQ ID NO:741);
ATAAAGCTTG	(SEQ ID NO:742);	TAAAGCTTGC	(SEQ ID NO:743);
AAAGCTTGCC	(SEQ ID NO:744);	AAGCTTGCCCT	(SEQ ID NO:745);
AGCTTGCCCT	(SEQ ID NO:746);	GCTTGCCCTTG	(SEQ ID NO:747);
CTTGCCCTTGA	(SEQ ID NO:748);	TTGCCCTTGAG	(SEQ ID NO:749);

- 223 -

TGCCTTGAGT	(SEQ ID NO:750);	GCCTTGAGTG	(SEQ ID NO:751);
CCTTGAGTGC	(SEQ ID NO:752);	CTTGAGTGCT	(SEQ ID NO:753);
TTGAGTGCTT	(SEQ ID NO:754);	TGAGTGCTTC	(SEQ ID NO:755);
GAGTGCTTCA	(SEQ ID NO:756);	AGTGCTTCAA	(SEQ ID NO:757);
GTGCTTCAAG	(SEQ ID NO:758);	TGCTTCAAGT	(SEQ ID NO:759);
GCTTCAAGTA	(SEQ ID NO:760);	CTTCAAGTAG	(SEQ ID NO:761);
TTCAAGTAGT	(SEQ ID NO:762);	TCAAGTAGTG	(SEQ ID NO:763);
CAAGTAGTGT	(SEQ ID NO:764);	AAGTAGTGTG	(SEQ ID NO:765);
AGTAGTGTGT	(SEQ ID NO:766);	GTAGTGTGTG	(SEQ ID NO:767);
TAGTGTGTGC	(SEQ ID NO:768);	AGTGTGTGCC	(SEQ ID NO:769);
GTGTGTGCCC	(SEQ ID NO:770);	TGTGTGCCCCG	(SEQ ID NO:771);
GTGTGCCCCGT	(SEQ ID NO:772);	TGTGCCCCGTC	(SEQ ID NO:773);
GTGCCCCGTCT	(SEQ ID NO:774);	TGCCCCGTCTG	(SEQ ID NO:775);
GCCCCGTCTGT	(SEQ ID NO:776);	CCCGTCTGTT	(SEQ ID NO:777);
CCGTCTGTTG	(SEQ ID NO:778);	CGTCTGTTGT	(SEQ ID NO:779);
GTCTGTTGTG	(SEQ ID NO:780);	TCTGTTGTGT	(SEQ ID NO:781);
CTGTTGTGTG	(SEQ ID NO:782);	TGTTGTGTGA	(SEQ ID NO:783);
GTTGTGTGAC	(SEQ ID NO:784);	TTGTGTGACT	(SEQ ID NO:785);
TGTGTGACTC	(SEQ ID NO:786);	GTGTGACTCT	(SEQ ID NO:787);
TGTGTGACTC	(SEQ ID NO:788);	GTGTGACTCT	(SEQ ID NO:789);
TGTGACTCTG	(SEQ ID NO:790);	GTGACTCTGG	(SEQ ID NO:791);
TGACTCTGGT	(SEQ ID NO:792);	GACTCTGGTA	(SEQ ID NO:793);
ACTCTGGTAA	(SEQ ID NO:794);	CTCTGGTAAC	(SEQ ID NO:795);
TCTGGTAACT	(SEQ ID NO:796);	CTGGTAACTA	(SEQ ID NO:797);
TGGTAACTAG	(SEQ ID NO:798);	GGTAACTAGA	(SEQ ID NO:799).

50. A method according to claim 24 wherein said HIV-1 isolate has the identifying characteristics of HIV-1 isolate ECACC V94101706 or HIV-1 isolate ECACC V941031169 or isolate C54 deposited at ECACC on 14 February, 1995 under Accession Number \_\_\_\_\_.

51. A method according to claim 20 wherein the HIV-1 isolate is reactive to antibodies to a glycoprotein of HIV-1, is capable of inducing an immune response of *gag*, *pol* and/or *env* and which is incapable of directing synthesis of a *nef* gene product or a full length *nef* gene product.

52. A method according to claim 51 wherein the glycoprotein is at least one of gp41-45, gp120 and/or gp160.
53. A method for obtaining a preparation of non-pathogenic HIV-1 from a biological sample, said method comprising co-culturing PBMCs from said biological sample from an individual putatively infected with said non-pathogenic HIV-1 with HIV-1 seronegative donor PBMCs depleted for CD8+ cells, harvesting the PBMCs and supernatant fluid every from about 5 to about 10 days and adding fresh medium with CD8+ depleted PBMCs with said fresh medium and isolating said virus from the supernatant fluid.
54. A method according to claim 53 wherein the seronegative PBMCs are pretreated with a cytokine for a time and under conditions sufficient to enhance HIV-1 replication.
55. A method according to claim 54 wherein the cytokine is M-CSF.
56. A method according to claim 53 or 54 or 55 wherein the HIV-1 seronegative donor PBMCs are first stimulated with PHA and/or IL-2.
57. A method according to claim 53 wherein the PBMCs from a putative infected individual are also CD8+ depleted prior to coculturing with seronegative donor PBMCs.
58. A method according to claim 53 or 57 wherein the cocultured cells are subjected to UV irradiation.
59. A method for obtaining a preparation of non-pathogenic HIV-1 from a biological sample, said method comprising co-culturing monocytes from said biological sample from an individual putatively infected with said non-pathogenic HIV-1 with HIV-1 seronegative donor PBMCs depleted for CD8+ cells, harvesting the monocytes and PBMCs and supernatant fluid every from about 5 to about 10 days and adding fresh medium with CD8+ depleted PBMCs with said fresh medium and isolating said virus from the supernatant fluid.

- 225 -

60. A method according to claim 59 wherein the seronegative PBMCs are pretreated with a cytokine for a time and under conditions sufficient to enhance HIV-1 replication.
61. A method according to claim 60 wherein the cytokine is M-CSF.
62. A method according to claim 59 or 60 or 61 wherein the HIV-1 seronegative donor PBMCs are first stimulated with PHA and/or IL-2.
63. A method according to claim 59 wherein the monocytes from a putative infected individual are also CD8+ depleted prior to coculturing with seronegative donor PBMCs.
64. A method according to claim 59 or 63 wherein the cocultured cells are subjected to UV irradiation.
65. A method for identifying or screening for compounds capable of reducing or otherwise interfering with HIV-1 replication, said method comprising contacting a compound to be tested with a cell or cell extract containing or capable of containing a *nef* gene product fused to a reporter molecule capable of giving an identifiable signal and screening for a compound which inhibits said signal.
66. A method according to claim 65 wherein the compound is an antibody to *nef* gene product or a part thereof.
67. A method according to claim 65 wherein the compound is a DNA targeting agent and inhibits transcription of the *nef* gene.
68. A compound capable of inhibiting *nef* gene product function or *nef* gene transcription.
69. A compound capable of inhibiting *nef* gene product function or *nef* gene transcription identified according to the method of claim 65 or 66 or 67.

70. A therapeutic composition useful for inhibiting or reducing productive infection by a pathogenic strain of HIV-1 and/or for vaccinating an individual against the development of AIDS or AIDS-related diseases, said composition comprising a non-pathogenic strain of HIV-1 according to any one of claims 1 to 16 and one or more pharmaceutical acceptable carriers and/or diluents.

71. A therapeutic composition according to claim 70 wherein said HIV-1 isolate comprises genetic material that directs expression of antisense or ribozyme nucleotide sequences which inhibit production of one or more proteins encoded by a pathogenic strain of HIV-1.

72. A viral isolate which:

- (i) is genetically or immunologically related to a pathogenic strain of HIV-1;
- (ii) is substantially non-pathogenic in human subjects;
- (iii) comprises a first nucleotide sequence constituting its genome which is capable of hybridising under medium stringency conditions to SEQ ID NO: 1 or a complementary form thereof; and
- (iv) comprises a second nucleotide sequence within said first nucleotide sequence and which second nucleotide sequence directs expression of a mRNA molecule capable of inhibiting, reducing or otherwise down-regulating translation of a protein or polypeptide encoded by a pathogenic strain of HIV-1 or inhibit, reduce or otherwise down regulate operation of a non-protein encoding a region of a pathogenic strain of HIV-1.

73. A viral isolate according to claim 72 wherein the targeted protein or polypeptide is selected from the list consisting of *gag* or *pol*.

74. A viral isolate according to claim 72 wherein the targeted protein or polypeptide is selected from the list consisting of *env* or *tat*.

75. A viral isolate according to claim 72 wherein the targeted protein or polypeptide is selected from the list consisting of *rev* or *vpu*.



- 227 -

76. A viral isolate according to claim 72 wherein the targeted protein or polypeptide is selected from the list consisting of *vpr*, *vif* or *nef*.

77. A viral isolate according to claim 72 wherein the targeted protein is *nef*.

78. A method for determining the pathogenicity of an HIV-1 strain after said HIV-1 strain infects cells of an individual, said method comprising determining the presence of a deletion mutation in the genome of said HIV-1 wherein said deletion mutation results in said genome being unable to synthesise a polypeptide or protein from a pathogenic strain of HIV-1 or directing the synthesis of a truncated form of said polypeptide or protein wherein the presence of a such a mutation is indicative of the presence of a non-pathogenic strain of HIV-1.

79. The method according to claim 78 wherein said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides within the region from nucleotide 8787 to nucleotide 9709 using the nucleotide numbering of HIV-1 strain NL4-3.

80. A strain of HIV-1 according to claim 79 wherein said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

nucleotide	(i)	8830-8862;
	(ii)	9009-9035;
	(iii)	9019-9029; and
	(iv)	9033-9049.

81. A strain of HIV-1 according to claim 79 wherein said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

nucleotide	(v)	9281-9371;
	(vi)	9281-9362;
	(vii)	9105-9224; and
	(viii)	9271-9370.

- 228 -

82. A strain of HIV-1 according to claim 79 wherein said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

nucleotide	(ix)	8882-8928;
	(x)	8850-9006;
	(xi)	8792-9041; and
	(xii)	9112-9204.

83. A strain of HIV-1 according to claim 79 wherein said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

nucleotide	(xiii)	9105-9224;
	(xiv)	9389-9395; and
	(xv)	9281-9366.

84. A strain of HIV-1 according to claim 79 wherein said deletion encompasses at least one of the following decanucleotides from the *nef* gene of HIV-1 strain NL4-3:

ATGGGTGGCA (SEQ ID NO:2);	TGGGTGGCAA (SEQ ID NO:3);
GGGTGGCAAG (SEQ ID NO:4);	GGTGGCAAGT (SEQ ID NO:5);
GTGGCAAGTG (SEQ ID NO:6);	TGGCAAGTGG (SEQ ID NO:7);
GGCAAGTGGT (SEQ ID NO:8);	GCAAGTGGTC (SEQ ID NO:9);
CAAGTGGTCA (SEQ ID NO:10);	AAGTGGTCAA (SEQ ID NO:11);
AGTGGTCAAA (SEQ ID NO:12);	GTGGTCAAAA (SEQ ID NO:13);
TGGTCAAAAA (SEQ ID NO:14);	GGTCAAAAAG (SEQ ID NO:15);
GTCAAAAAGT (SEQ ID NO:16);	TCAAAAAGTA (SEQ ID NO:17);
CAAAAAGTAG (SEQ ID NO:18);	AAAAAGTAGT (SEQ ID NO:19);
AAAAGTAGTG (SEQ ID NO:20);	AAAGTAGTGT (SEQ ID NO:21);
AAGTAGTGTG (SEQ ID NO:22);	AGTAGTGTGA (SEQ ID NO:23);
GTAGTGTGAT (SEQ ID NO:24);	TAGTGTGATT (SEQ ID NO:25);
AGTGTGATTG (SEQ ID NO:26);	GTGTGATTGG (SEQ ID NO:27);
TGTGATTGGA (SEQ ID NO:28);	GTGATTGGAT (SEQ ID NO:29);
TGATTGGATG (SEQ ID NO:30);	GATTGGATGG (SEQ ID NO:31);
ATTGGATGGC (SEQ ID NO:32);	TTGGATGGCC (SEQ ID NO:33);
TGGATGGCCT (SEQ ID NO:34);	GGATGGCCTG (SEQ ID NO:35);
GATGGCCTGC (SEQ ID NO:36);	ATGGCCTGCT (SEQ ID NO:37);

TGGCCTGCTG (SEQ ID NO:38);	GGCCTGCTGT (SEQ ID NO:39);
GCCTGCTGTA (SEQ ID NO:40);	CCTGCTGTAA (SEQ ID NO:41);
CTGCTGTAAG (SEQ ID NO:42);	TGCTGTAAGG (SEQ ID NO:43);
GCTGTAAGGG (SEQ ID NO:44);	CTGTAAGGGA (SEQ ID NO:45);
TGTAAGGGAA (SEQ ID NO:46);	GTAAGGGAAA (SEQ ID NO:47);
TAAGGGAAAG (SEQ ID NO:48);	AAGGGAAAGA (SEQ ID NO:49);
AGGGAAAGAA (SEQ ID NO:50);	GGGAAAGAAT (SEQ ID NO:51);
GGAAAGAATG (SEQ ID NO:52);	GAAAGAATGA (SEQ ID NO:53);
AAAGAATGAG (SEQ ID NO:54);	AAGAATGAGA (SEQ ID NO:55);
AGAATGAGAC (SEQ ID NO:56);	GAATGAGACG (SEQ ID NO:57);
AATGAGACGA (SEQ ID NO:58);	ATGAGACGAG (SEQ ID NO:59);
TGAGACGAGC (SEQ ID NO:60);	GAGACGAGCT (SEQ ID NO:61);
AGACGAGCTG (SEQ ID NO:62);	GACGAGCTGA (SEQ ID NO:63);
ACGAGCTGAG (SEQ ID NO:64);	CGAGCTGAGC (SEQ ID NO:65);
GAGCTGAGCC (SEQ ID NO:66);	AGCTGAGCCA (SEQ ID NO:67);
GCTGAGCCAG (SEQ ID NO:68);	CTGAGCCAGC (SEQ ID NO:69);
TGAGCCAGCA (SEQ ID NO:70);	GAGCCAGCAG (SEQ ID NO:71);
AGCCAGCAGC (SEQ ID NO:72);	GCCAGCAGCA (SEQ ID NO:73);
CCAGCAGCAG (SEQ ID NO:74);	CAGCAGCAGA (SEQ ID NO:75);
AGCAGCAGAT (SEQ ID NO:76);	GCAGCAGATG (SEQ ID NO:77);
CAGCAGATGG (SEQ ID NO:78);	AGCAGATGGG (SEQ ID NO:79);
GCAGATGGGG (SEQ ID NO:80);	CAGATGGGGT (SEQ ID NO:81);
AGATGGGGTG (SEQ ID NO:82);	GATGGGGTGG (SEQ ID NO:83);
ATGGGGTGGG (SEQ ID NO:84);	TGGGGTGGGA (SEQ ID NO:85);
GGGGTGGGAG (SEQ ID NO:86);	GGGTGGGAGC (SEQ ID NO:87);
GGTGGGAGCA (SEQ ID NO:88);	GTGGGAGCAG (SEQ ID NO:89);
TGGGAGCAGT (SEQ ID NO:90);	GGGAGCAGTA (SEQ ID NO:91);
GGAGCAGTAT (SEQ ID NO:92);	GAGCAGTATC (SEQ ID NO:93);
AGCAGTATCT (SEQ ID NO:94);	GCAGTATCTC (SEQ ID NO:95);
CAGTATCTCG (SEQ ID NO:96);	AGTATCTCGA (SEQ ID NO:97);
GTATCTCGAG (SEQ ID NO:98);	TATCTCGAGA (SEQ ID NO:99);
ATCTCGAGAC (SEQ ID NO:100);	TCTCGAGACC (SEQ ID NO:101);
CTCGAGACCT (SEQ ID NO:102);	TCGAGACCTA (SEQ ID NO:103);
CGAGACCTAG (SEQ ID NO:104);	GAGACCTAGA (SEQ ID NO:105);
AGACCTAGAA (SEQ ID NO:106);	GACCTAGAAA (SEQ ID NO:107);
ACCTAGAAAA (SEQ ID NO:108);	CCTAGAAAAA (SEQ ID NO:109);
CTAGAAAAAC (SEQ ID NO:110);	TAGAAAAACA (SEQ ID NO:111);
AGAAAAACAT (SEQ ID NO:112);	GAAAAACATG (SEQ ID NO:113);

- 230 -

AAAAACATGG (SEQ ID NO:114);	AAAACATGGA (SEQ ID NO:115);
AAACATGGAG (SEQ ID NO:116);	AACATGGAGC (SEQ ID NO:117);
ACATGGAGCA (SEQ ID NO:118);	CATGGAGCAA (SEQ ID NO:119);
ATGGAGCAAT (SEQ ID NO:120);	TGGAGCAATC (SEQ ID NO:121);
GGAGCAATCA (SEQ ID NO:122);	GAGCAATCAC (SEQ ID NO:123);
AGCAATCACA (SEQ ID NO:124);	GCAATCACAA (SEQ ID NO:125);
CAATCACAAG (SEQ ID NO:126);	AATCACAAGT (SEQ ID NO:127);
ATCACAAGTA (SEQ ID NO:128);	TCACAAGTAG (SEQ ID NO:129);
CACAAGTAGC (SEQ ID NO:130);	ACAAGTAGCA (SEQ ID NO:131);
CAAGTAGCAA (SEQ ID NO:132);	AAGTAGCAAT (SEQ ID NO:133);
AGTAGCAATA (SEQ ID NO:134);	GTAGCAATAC (SEQ ID NO:135);
TAGCAATACA (SEQ ID NO:136);	AGCAATACAG (SEQ ID NO:137);
GCAATACAGC (SEQ ID NO:138);	CAATACAGCA (SEQ ID NO:139);
AATACAGCAG (SEQ ID NO:140);	ATACAGCAGC (SEQ ID NO:141);
TACAGCAGCT (SEQ ID NO:142);	ACAGCAGCTA (SEQ ID NO:143);
CAGCAGCTAA (SEQ ID NO:144);	AGCAGCTAAC (SEQ ID NO:145);
GCAGCTAACA (SEQ ID NO:146);	CAGCTAACAA (SEQ ID NO:147);
AGCTAACAAT (SEQ ID NO:148);	GCTAACAATG (SEQ ID NO:149);
CTAACAATGC (SEQ ID NO:150);	TAACAATGCT (SEQ ID NO:151);
AACAATGCTG (SEQ ID NO:152);	ACAATGCTGC (SEQ ID NO:153);
CAATGCTGCT (SEQ ID NO:154);	AATGCTGCTT (SEQ ID NO:155);
ATGCTGCTTG (SEQ ID NO:156);	TGCTGCTTGT (SEQ ID NO:157);
GCTGCTTGTG (SEQ ID NO:158);	CTGCTTGTGC (SEQ ID NO:159);
TGCTTGTGCC (SEQ ID NO:160);	GCTTGTGCCT (SEQ ID NO:161);
CTTGTGCCTG (SEQ ID NO:162);	TTGTGCCTGG (SEQ ID NO:163);
TGTGCCTGGC (SEQ ID NO:164);	GTGCCTGGCT (SEQ ID NO:165);
TGCCTGGCTA (SEQ ID NO:166);	GCCTGGCTAG (SEQ ID NO:167);
CCTGGCTAGA (SEQ ID NO:168);	CTGGCTAGAA (SEQ ID NO:169);
TGGCTAGAAG (SEQ ID NO:170);	GGCTAGAAGC (SEQ ID NO:171);
GCTAGAAGCA (SEQ ID NO:172);	CTAGAAGCAC (SEQ ID NO:173);
TAGAAGCACA (SEQ ID NO:174);	AGAAGCACAA (SEQ ID NO:175);
GAAGCACAAG (SEQ ID NO:176);	AAGCACAAGA (SEQ ID NO:177);
AGCACAAGAG (SEQ ID NO:178);	GCACAAGAGG (SEQ ID NO:179);
CACAAGAGGA (SEQ ID NO:180);	ACAAGAGGAG (SEQ ID NO:181);
CAAGAGGAGG (SEQ ID NO:182);	AAGAGGAGGA (SEQ ID NO:183);
AGAGGAGGAA (SEQ ID NO:184);	GAGGAGGAAG (SEQ ID NO:185);
AGGAGGAAGA (SEQ ID NO:186);	GGAGGAAGAG (SEQ ID NO:187);
GAGGAAGAGG (SEQ ID NO:188);	AGGAAGAGGT (SEQ ID NO:189);

- 231 -

GGAAGAGGTG (SEQ ID NO:190);	GAAGAGGTGG (SEQ ID NO:191);
AAGAGGTGGG (SEQ ID NO:192);	AGAGGTGGGT (SEQ ID NO:193);
GAGGTGGGTT (SEQ ID NO:194);	AGGTGGGTTT (SEQ ID NO:195);
GGTGGGTTTT (SEQ ID NO:196);	GTGGGTTTTT (SEQ ID NO:197);
TGGGTTTTCC (SEQ ID NO:198);	GGGTTTTCCA (SEQ ID NO:199);
GGTTTTCCAG (SEQ ID NO:200);	GTTTTCCAGT (SEQ ID NO:201);
TTTTCCAGTC (SEQ ID NO:202);	TTTCCAGTCA (SEQ ID NO:203);
TTCCAGTCAC (SEQ ID NO:204);	TCCAGTCACA (SEQ ID NO:205);
CCAGTCACAC (SEQ ID NO:206);	CAGTCACACC (SEQ ID NO:207);
AGTCACACCT (SEQ ID NO:208);	GTCACACCTC (SEQ ID NO:209);
TCACACCTCA (SEQ ID NO:210);	CACACCTCAG (SEQ ID NO:211);
ACACCTCAGG (SEQ ID NO:212);	CACCTCAGGT (SEQ ID NO:213);
ACCTCAGGTA (SEQ ID NO:214);	CCTCAGGTAC (SEQ ID NO:215);
CTCAGGTACC (SEQ ID NO:216);	TCAGGTACCT (SEQ ID NO:217);
CAGGTACCTT (SEQ ID NO:218);	AGGTACCTTT (SEQ ID NO:219);
GGTACCTTTA (SEQ ID NO:220);	GTACCTTTAA (SEQ ID NO:221);
TACCTTTAAG (SEQ ID NO:222);	ACCTTTAAGA (SEQ ID NO:223);
CCTTTAAGAC (SEQ ID NO:224);	CTTTAAGACC (SEQ ID NO:225);
TTTAAGACCA (SEQ ID NO:226);	TTAAGACCAA (SEQ ID NO:227);
TAAGACCAAT (SEQ ID NO:228);	AAGACCAATG (SEQ ID NO:229);
AGACCAATGA (SEQ ID NO:230);	GACCAATGAC (SEQ ID NO:231);
ACCAATGACT (SEQ ID NO:232);	CCAATGACTT (SEQ ID NO:233);
CAATGACTTA (SEQ ID NO:234);	AATGACTTAC (SEQ ID NO:235);
ATGACTTACA (SEQ ID NO:236);	TGACTTACAA (SEQ ID NO:237);
GACTTACAAG (SEQ ID NO:238);	ACTTACAAGG (SEQ ID NO:239);
CTTACAAGGC (SEQ ID NO:240);	TTACAAGGCA (SEQ ID NO:241);
TACAAGGCAG (SEQ ID NO:242);	ACAAGGCAGC (SEQ ID NO:243);
CAAGGCAGCT (SEQ ID NO:244);	AAGGCAGCTG (SEQ ID NO:245);
AGGCAGCTGT (SEQ ID NO:246);	GGCAGCTGTA (SEQ ID NO:247);
GCAGCTGTAG (SEQ ID NO:248);	CAGCTGTAGA (SEQ ID NO:249);
AGCTGTAGAT (SEQ ID NO:250);	GCTGTAGATC (SEQ ID NO:251);
CTGTAGATCT (SEQ ID NO:252);	TGTAGATCTT (SEQ ID NO:253);
GTAGATCTTA (SEQ ID NO:254);	TAGATCTTAG (SEQ ID NO:255);
AGATCTTAGC (SEQ ID NO:256);	GATCTTAGCC (SEQ ID NO:257);
ATCTTAGCCA (SEQ ID NO:258);	TCTTAGCCAC (SEQ ID NO:259);
CTTAGCCACT (SEQ ID NO:260);	TTAGCCACTT (SEQ ID NO:261);
TAGCCACTTT (SEQ ID NO:262);	AGCCACTTTT (SEQ ID NO:263);
GCCACTTTTT (SEQ ID NO:264);	CCACTTTTTA (SEQ ID NO:265);

- 232 -

CACTTTTTTAA (SEQ ID NO:266);	ACTTTTTTAAA (SEQ ID NO:267);
CTTTTTTAAAA (SEQ ID NO:268);	TTTTTTAAAAG (SEQ ID NO:269);
TTTTTTAAAGA (SEQ ID NO:270);	TTTAAAGAA (SEQ ID NO:271);
TTAAAGAAAA (SEQ ID NO:272);	TAAAGAAAA (SEQ ID NO:273);
AAAAGAAAAG (SEQ ID NO:274);	AAAGAAAAGG (SEQ ID NO:275);
AAGAAAAGGG (SEQ ID NO:276);	AGAAAAGGGG (SEQ ID NO:277);
GAAAAGGGGG (SEQ ID NO:278);	AAAAGGGGGG (SEQ ID NO:279);
AAAGGGGGGA (SEQ ID NO:280);	AAGGGGGGAC (SEQ ID NO:281);
AGGGGGGACT (SEQ ID NO:282);	GGGGGGACTG (SEQ ID NO:283);
GGGGGACTGG (SEQ ID NO:284);	GGGGACTGGA (SEQ ID NO:285);
GGGACTGGAA (SEQ ID NO:286);	GGACTGGAAG (SEQ ID NO:287);
GACTGGAAGG (SEQ ID NO:288);	ACTGGAAGGG (SEQ ID NO:289);
CTGGAAGGGC (SEQ ID NO:290);	TGGAAGGGCT (SEQ ID NO:291);
GGAAGGGCTA (SEQ ID NO:292);	GAAGGGCTAA (SEQ ID NO:293);
AAGGGCTAAT (SEQ ID NO:294);	AGGGCTAATT (SEQ ID NO:295);
GGGCTAATTC (SEQ ID NO:296);	GGCTAATTCA (SEQ ID NO:297);
GCTAATTCAC (SEQ ID NO:298);	CTAATTCACT (SEQ ID NO:299);
TAATTCACTC (SEQ ID NO:300);	AATTCACTCC (SEQ ID NO:301);
ATTCACTCCC (SEQ ID NO:302);	TTCACTCCCA (SEQ ID NO:303);
TCACTCCCAA (SEQ ID NO:304);	CACTCCCAAA (SEQ ID NO:305);
ACTCCCAAAG (SEQ ID NO:306);	CTCCCAAAGA (SEQ ID NO:307);
TCCCAAAGAA (SEQ ID NO:308);	CCCAAAGAAG (SEQ ID NO:309);
CCAAAGAAGA (SEQ ID NO:310);	CAAAGAAGAC (SEQ ID NO:311);
AAAGAAGACA (SEQ ID NO:312);	AAGAAGACAA (SEQ ID NO:313);
AGAAGACAAG (SEQ ID NO:314);	GAAGACAAGA (SEQ ID NO:315);
AAGACAAGAT (SEQ ID NO:316);	AGACAAGATA (SEQ ID NO:317);
GACAAGATAT (SEQ ID NO:318);	ACAAGATATC (SEQ ID NO:319);
CAAGATATCC (SEQ ID NO:320);	AAGATATCCT (SEQ ID NO:321);
AGATATCCTT (SEQ ID NO:322);	GATATCCTTG (SEQ ID NO:323);
ATATCCTTGA (SEQ ID NO:324);	TATCCTTGAT (SEQ ID NO:325);
ATCCTTGATC (SEQ ID NO:326);	TCCTTGATCT (SEQ ID NO:327);
CCTTGATCTG (SEQ ID NO:328);	CTTGATCTGT (SEQ ID NO:329);
TTGATCTGTG (SEQ ID NO:330);	TGATCTGTGG (SEQ ID NO:331);
GATCTGTGGA (SEQ ID NO:332);	ATCTGTGGAT (SEQ ID NO:333);
TCTGTGGATC (SEQ ID NO:334);	CTGTGGATCT (SEQ ID NO:335);
TGTGGATCTA (SEQ ID NO:336);	GTGGATCTAC (SEQ ID NO:337);
TGGATCTACC (SEQ ID NO:338);	GGATCTACCA (SEQ ID NO:339);
GATCTACCAC (SEQ ID NO:340);	ATCTACCACA (SEQ ID NO:341);

- 233 -

TCTACCACAC (SEQ ID NO:342);	CTACCACACA (SEQ ID NO:343);
TACCACACAC (SEQ ID NO:344);	ACCACACACA (SEQ ID NO:345);
CCACACACAA (SEQ ID NO:346);	CACACACAAG (SEQ ID NO:347);
ACACACAAGG (SEQ ID NO:348);	CACACAAGGC (SEQ ID NO:349);
ACACAAGGCT (SEQ ID NO:350);	CACAAGGCTA (SEQ ID NO:351);
ACAAGGCTAC (SEQ ID NO:352);	CAAGGCTACT (SEQ ID NO:353);
AAGGCTACTT (SEQ ID NO:354);	AGGCTACTTC (SEQ ID NO:355);
GGCTACTTCC (SEQ ID NO:356);	GCTACTTCCC (SEQ ID NO:357);
CTACTTCCCT (SEQ ID NO:358);	TACTTCCCTG (SEQ ID NO:359);
ACTTCCCTGA (SEQ ID NO:360);	CTTCCCTGAT (SEQ ID NO:361);
TTCCCTGATT (SEQ ID NO:362);	TCCCTGATTG (SEQ ID NO:363);
CCCTGATTGG (SEQ ID NO:364);	CCTGATTGGC (SEQ ID NO:365);
CTGATTGGCA (SEQ ID NO:366);	TGATTGGCAG (SEQ ID NO:367);
GATTGGCAGA (SEQ ID NO:368);	ATTGGCAGAA (SEQ ID NO:369);
TTGGCAGAAC (SEQ ID NO:370);	TGGCAGAACT (SEQ ID NO:371);
GGCAGAACTA (SEQ ID NO:372);	GCAGAACTAC (SEQ ID NO:373);
CAGAACTACA (SEQ ID NO:374);	AGAACTACAC (SEQ ID NO:375);
GAACTACACA (SEQ ID NO:376);	AACTACACAC (SEQ ID NO:377);
ACTACACACC (SEQ ID NO:378);	CTACACACCA (SEQ ID NO:379);
TACACACCAG (SEQ ID NO:380);	ACACACCAGG (SEQ ID NO:381);
CACACCAGGG (SEQ ID NO:382);	ACACCAGGGC (SEQ ID NO:383);
CACCAGGGCC (SEQ ID NO:384);	ACCAGGGCCA (SEQ ID NO:385);
CCAGGGCCAG (SEQ ID NO:386);	CAGGGCCAGG (SEQ ID NO:387);
AGGGCCAGGG (SEQ ID NO:388);	GGGCCAGGGG (SEQ ID NO:389);
GGCCAGGGGT (SEQ ID NO:390);	GCCAGGGGTC (SEQ ID NO:391);
CCAGGGGTCA (SEQ ID NO:392);	CAGGGGTCAG (SEQ ID NO:393);
AGGGGTCAGA (SEQ ID NO:394);	GGGGTCAGAT (SEQ ID NO:395);
GGGTCAGATA (SEQ ID NO:396);	GGTCAGATAT (SEQ ID NO:397);
GTCAGATATC (SEQ ID NO:398);	TCAGATATCC (SEQ ID NO:399);
CAGATATCCA (SEQ ID NO:400);	AGATATCCAC (SEQ ID NO:401);
GATATCCACT (SEQ ID NO:402);	ATATCCACTG (SEQ ID NO:403);
TATCCACTGA (SEQ ID NO:404);	ATCCACTGAC (SEQ ID NO:405);
TCCACTGACC (SEQ ID NO:406);	CCACTGACCT (SEQ ID NO:407);
CACTGACCTT (SEQ ID NO:408);	ACTGACCTTT (SEQ ID NO:409);
CTGACCTTTG (SEQ ID NO:410);	TGACCTTTGG (SEQ ID NO:411);
GACCTTTGGA (SEQ ID NO:412);	ACCTTTGGAT (SEQ ID NO:413);
CCTTTGGATG (SEQ ID NO:414);	CTTTGGATGG (SEQ ID NO:415);
TTTGGATGGT (SEQ ID NO:416);	TTGGATGGTG (SEQ ID NO:417);

- 234 -

TGGATGGTGC (SEQ ID NO:418);	GGATGGTGCT (SEQ ID NO:419);
GATGGTGCTA (SEQ ID NO:420);	ATGGTGCTAC (SEQ ID NO:421);
TGGTGCTACA (SEQ ID NO:422);	GGTGCTACAA (SEQ ID NO:423);
GTGCTACAAG (SEQ ID NO:424);	TGCTACAAGC (SEQ ID NO:425);
GCTACAAGCT (SEQ ID NO:426);	CTACAAGCTA (SEQ ID NO:427);
TACAAGCTAG (SEQ ID NO:428);	ACAAGCTAGT (SEQ ID NO:429);
CAAGCTAGTA (SEQ ID NO:430);	AAGCTAGTAC (SEQ ID NO:431);
AGCTAGTACC (SEQ ID NO:432);	GCTAGTACCA (SEQ ID NO:433);
CTAGTACCAG (SEQ ID NO:434);	TAGTACCAGT (SEQ ID NO:435);
AGTACCAGTT (SEQ ID NO:436);	GTACCAGTTG (SEQ ID NO:437);
TACCAGTTGA (SEQ ID NO:438);	ACCAGTTGAG (SEQ ID NO:439);
CCAGTTGAGC (SEQ ID NO:440);	CAGTTGAGCC (SEQ ID NO:441);
AGTTGAGCCA (SEQ ID NO:442);	GTTGAGCCAG (SEQ ID NO:443);
TTGAGCCAGA (SEQ ID NO:444);	TGAGCCAGAT (SEQ ID NO:445);
GAGCCAGATA (SEQ ID NO:446);	AGCCAGATAA (SEQ ID NO:447);
GCCAGATAAG (SEQ ID NO:448);	CCAGATAAGG (SEQ ID NO:449);
CAGATAAGGT (SEQ ID NO:450);	AGATAAGGTA (SEQ ID NO:451);
GATAAGGTAG (SEQ ID NO:452);	ATAAGGTAGA (SEQ ID NO:453);
TAAGGTAGAA (SEQ ID NO:454);	AAGGTAGAAG (SEQ ID NO:455);
AGGTAGAAGA (SEQ ID NO:456);	GGTAGAAGAG (SEQ ID NO:457);
GTAGAAGAGG (SEQ ID NO:458);	TAGAAGAGGC (SEQ ID NO:459);
AGAAGAGGCC (SEQ ID NO:460);	GAAGAGGCCA (SEQ ID NO:461);
AAGAGGCCAA (SEQ ID NO:462);	AGAGGCCAAT (SEQ ID NO:463);
GAGGCCAATA (SEQ ID NO:464);	AGGCCAATAA (SEQ ID NO:465);
GGCCAATAAA (SEQ ID NO:466);	GCCAATAAAG (SEQ ID NO:467);
CCAATAAAGG (SEQ ID NO:468);	CAATAAAGGA (SEQ ID NO:469);
AATAAAGGAG (SEQ ID NO:470);	ATAAAGGAGA (SEQ ID NO:471);
TAAAGGAGAG (SEQ ID NO:472);	AAAGGAGAGA (SEQ ID NO:473);
AAGGAGAGAA (SEQ ID NO:474);	AGGAGAGAAC (SEQ ID NO:475);
GGAGAGAACA (SEQ ID NO:476);	GAGAGAACAC (SEQ ID NO:477);
AGAGAACACC (SEQ ID NO:478);	GAGAACACCA (SEQ ID NO:479);
AGAACACCAG (SEQ ID NO:480);	GAACACCAGC (SEQ ID NO:481);
AACACCAGCT (SEQ ID NO:482);	ACACCAGCTT (SEQ ID NO:483);
CACCAGCTTG (SEQ ID NO:484);	ACCAGCTTGT (SEQ ID NO:485);
CCAGCTTGTT (SEQ ID NO:486);	CAGCTTGTTA (SEQ ID NO:487);
AGCTTGTTAC (SEQ ID NO:488);	GCTTGTTACA (SEQ ID NO:489);
CTTGTTACAC (SEQ ID NO:490);	TTGTTACACC (SEQ ID NO:491);
TGTTACACCC (SEQ ID NO:492);	GTTACACCCT (SEQ ID NO:493);



TTACACCCTG (SEQ ID NO:494);	TACACCCTGT (SEQ ID NO:495);
ACACCCTGTG (SEQ ID NO:496);	CACCCTGTGA (SEQ ID NO:497);
ACCCTGTGAG (SEQ ID NO:498);	CCCTGTGAGC (SEQ ID NO:499);
CCTGTGAGCC (SEQ ID NO:500);	CTGTGAGCCT (SEQ ID NO:501);
TGTGAGCCTG (SEQ ID NO:502);	GTGAGCCTGC (SEQ ID NO:503);
TGAGCCTGCA (SEQ ID NO:504);	GAGCCTGCAT (SEQ ID NO:505);
AGCCTGCATG (SEQ ID NO:506);	GCCTGCATGG (SEQ ID NO:507);
CCTGCATGGA (SEQ ID NO:508);	CTGCATGGAA (SEQ ID NO:509);
TGCATGGAAT (SEQ ID NO:510);	GCATGGAATG (SEQ ID NO:511);
CATGGAATGG (SEQ ID NO:512);	ATGGAATGGA (SEQ ID NO:513);
TGGAATGGAT (SEQ ID NO:514);	GGAATGGATG (SEQ ID NO:515);
GAATGGATGA (SEQ ID NO:516);	AATGGATGAC (SEQ ID NO:517);
ATGGATGACC (SEQ ID NO:518);	TGGATGACCC (SEQ ID NO:519);
GGATGACCCT (SEQ ID NO:520);	GATGACCCTG (SEQ ID NO:521);
ATGACCCTGA (SEQ ID NO:522);	TGACCCTGAG (SEQ ID NO:523);
GACCCTGAGA (SEQ ID NO:524);	ACCCTGAGAG (SEQ ID NO:525);
CCCTGAGAGA (SEQ ID NO:526);	CCTGAGAGAG (SEQ ID NO:527);
CTGAGAGAGA (SEQ ID NO:528);	TGAGAGAGAA (SEQ ID NO:529);
GAGAGAGAAG (SEQ ID NO:530);	AGAGAGAAGT (SEQ ID NO:531);
GAGAGAAGTG (SEQ ID NO:532);	AGAGAAGTGT (SEQ ID NO:533);
GAGAAGTGTT (SEQ ID NO:534);	AGAAGTGTTA (SEQ ID NO:535);
GAAGTGTTAG (SEQ ID NO:536);	AAGTGTTAGA (SEQ ID NO:537);
AGTGTTAGAG (SEQ ID NO:538);	GTGTTAGAGT (SEQ ID NO:539);
TGTTAGAGTG (SEQ ID NO:540);	GTTAGAGTGG (SEQ ID NO:541);
TTAGAGTGGA (SEQ ID NO:542);	TAGAGTGGAG (SEQ ID NO:543);
AGAGTGGAGG (SEQ ID NO:544);	GAGTGGAGGT (SEQ ID NO:545);
AGTGGAGGTT (SEQ ID NO:546);	GTGGAGGTTT (SEQ ID NO:547);
TGGAGGTTTG (SEQ ID NO:548);	GGAGGTTTGA (SEQ ID NO:549);
GAGGTTTGAC (SEQ ID NO:550);	AGGTTTGACA (SEQ ID NO:551);
GGTTTGACAG (SEQ ID NO:552);	GTTTGACAGC (SEQ ID NO:553);
TTTGACAGCC (SEQ ID NO:554);	TTGACAGCCG (SEQ ID NO:555);
TGACAGCCGC (SEQ ID NO:556);	GACAGCCGCC (SEQ ID NO:557);
ACAGCCGCCT (SEQ ID NO:558);	CAGCCGCCTA (SEQ ID NO:559);
AGCCGCCTAG (SEQ ID NO:560);	GCCGCCTAGC (SEQ ID NO:561);
CCGCCTAGCA (SEQ ID NO:562);	CGCCTAGCAT (SEQ ID NO:563);
GCCTAGCATT (SEQ ID NO:564);	CCTAGCATTT (SEQ ID NO:565);
CTAGCATTTT (SEQ ID NO:566);	TAGCATTTCA (SEQ ID NO:567);
AGCATTTTCAT (SEQ ID NO:568);	GCATTTTCATC (SEQ ID NO:569);

- 236 -

CATTTTCATCA (SEQ ID NO:570);	ATTTTCATCAC (SEQ ID NO:571);
TTTCATCACG (SEQ ID NO:572);	TTCATCACGT (SEQ ID NO:573);
TCATCACGTG (SEQ ID NO:574);	CATCACGTGG (SEQ ID NO:575);
ATCACGTGGC (SEQ ID NO:576);	TCACGTGGCC (SEQ ID NO:577);
CACGTGGCCC (SEQ ID NO:578);	ACGTGGCCCCG (SEQ ID NO:579);
CGTGGCCCGA (SEQ ID NO:580);	GTGGCCCGAG (SEQ ID NO:581);
TGGCCCGAGA (SEQ ID NO:582);	GGCCCGAGAG (SEQ ID NO:583);
GCCCGAGAGC (SEQ ID NO:584);	CCCGAGAGCT (SEQ ID NO:585);
CCGAGAGCTG (SEQ ID NO:586);	CGAGAGCTGC (SEQ ID NO:587);
GAGAGCTGCA (SEQ ID NO:588);	AGAGCTGCAT (SEQ ID NO:589);
GAGCTGCATC (SEQ ID NO:590);	AGCTGCATCC (SEQ ID NO:591);
GCTGCATCCG (SEQ ID NO:592);	CTGCATCCGG (SEQ ID NO:593);
TGCATCCGGA (SEQ ID NO:594);	GCATCCGGAG (SEQ ID NO:595);
CATCCGGAGT (SEQ ID NO:596);	ATCCGGAGTA (SEQ ID NO:597);
TCCGGAGTAC (SEQ ID NO:598);	CCGGAGTACT (SEQ ID NO:599);
CGGAGTACTT (SEQ ID NO:600);	GGAGTACTTC (SEQ ID NO:601);
GAGTACTTCA (SEQ ID NO:602);	AGTACTTCAA (SEQ ID NO:603);
GTACTTCAAG (SEQ ID NO:604);	TACTTCAAGA (SEQ ID NO:605);
ACTTCAAGAA (SEQ ID NO:606);	CTTCAAGAAC (SEQ ID NO:607);
TTCAAGAACT (SEQ ID NO:608);	TCAAGAACTG (SEQ ID NO:609);
CAAGAACTGC (SEQ ID NO:610);	AAGAACTGCT (SEQ ID NO:611);
AGAACTGCTG (SEQ ID NO:612);	GAACTGCTGA (SEQ ID NO:613);

85. A strain of HIV-1 according to claim 79 wherein said deletion encompasses at least one of the following decanucleotides from the LTR region of HIV-1 strain NL4-3:

GCTTTTTGCC (SEQ ID NO:652);	CTTTTTGCCT (SEQ ID NO:653);
TTTTTGCCTG (SEQ ID NO:654);	TTTGCCTGT (SEQ ID NO:655);
TTTGCCTGTA (SEQ ID NO:656);	TTGCCTGTAC (SEQ ID NO:657);
TGCCTGTACT (SEQ ID NO:658);	GCCTGTACTG (SEQ ID NO:659);
CCTGTACTGG (SEQ ID NO:660);	CTGTACTGGG (SEQ ID NO:661);
TGTACTGGGT (SEQ ID NO:662);	GTACTGGGTC (SEQ ID NO:663);
TACTGGGTCT (SEQ ID NO:664);	ACTGGGTCTC (SEQ ID NO:665);
CTGGGTCTCT (SEQ ID NO:666);	TGGGTCTCTC (SEQ ID NO:667);
GGGTCTCTCT (SEQ ID NO:668);	GGTCTCTCTG (SEQ ID NO:669);
GTCTCTCTGG (SEQ ID NO:670);	TCTCTCTGGT (SEQ ID NO:671);
CTCTCTGGTT (SEQ ID NO:672);	TCTCTGGTTA (SEQ ID NO:673);
CTCTGGTTAG (SEQ ID NO:674);	TCTCTGGTTA (SEQ ID NO:675);

CTGGTTAGAC	(SEQ ID NO:676);	TGGTTAGACC	(SEQ ID NO:677);
GGTTAGACCA	(SEQ ID NO:678);	GTTAGACCAG	(SEQ ID NO:679);
TTAGACCAGA	(SEQ ID NO:680);	TAGACCAGAT	(SEQ ID NO:681);
AGACCAGATC	(SEQ ID NO:682);	GACCAGATCT	(SEQ ID NO:683);
ACCAGATCTG	(SEQ ID NO:684);	CCAGATCTGA	(SEQ ID NO:685);
CAGATCTGAG	(SEQ ID NO:686);	AGATCTGAGC	(SEQ ID NO:687);
GATCTGAGCC	(SEQ ID NO:688);	ATCTGAGCCT	(SEQ ID NO:689);
TCTGAGCCTG	(SEQ ID NO:690);	CTGAGCCTGG	(SEQ ID NO:691);
TGAGCCTGGG	(SEQ ID NO:692);	GAGCCTGGGA	(SEQ ID NO:693);
AGCCTGGGAG	(SEQ ID NO:694);	GCCTGGGAGC	(SEQ ID NO:695);
CCTGGGAGCT	(SEQ ID NO:696);	CTGGGAGCTC	(SEQ ID NO:697);
TGGGAGCTCT	(SEQ ID NO:698);	GGGAGCTCTC	(SEQ ID NO:699);
GGAGCTCTCT	(SEQ ID NO:700);	GAGCTCTCTG	(SEQ ID NO:701);
AGCTCTCTGG	(SEQ ID NO:702);	GCTCTCTGGC	(SEQ ID NO:703);
CTCTCTGGCT	(SEQ ID NO:704);	TCTCTGGCTA	(SEQ ID NO:705);
CTCTGGCTAA	(SEQ ID NO:706);	TCTGGCTAAC	(SEQ ID NO:707);
CTGGCTAACT	(SEQ ID NO:708);	TGGCTAACTA	(SEQ ID NO:709);
GGCTAACTAG	(SEQ ID NO:710);	GCTAACTAGG	(SEQ ID NO:711);
CTAACTAGGG	(SEQ ID NO:712);	TAACTAGGGA	(SEQ ID NO:713);
AACTAGGGAA	(SEQ ID NO:714);	ACTAGGGAAC	(SEQ ID NO:715);
CTAGGGAACC	(SEQ ID NO:716);	TAGGGAACCC	(SEQ ID NO:717);
AGGGAACCCA	(SEQ ID NO:718);	GGAACCCAC	(SEQ ID NO:719);
GGAACCCACT	(SEQ ID NO:720);	GAACCCACTG	(SEQ ID NO:721);
AACCCACTGC	(SEQ ID NO:722);	ACCCACTGCT	(SEQ ID NO:723);
CCCACTGCTT	(SEQ ID NO:724);	CCACTGCTTA	(SEQ ID NO:725);
CACTGCTTAA	(SEQ ID NO:726);	ACTGCTTAAG	(SEQ ID NO:727);
CTGCTTAAGC	(SEQ ID NO:728);	TGCTTAAGCC	(SEQ ID NO:729);
GCTTAAGCCT	(SEQ ID NO:730);	CTTAAGCCTC	(SEQ ID NO:731);
TTAAGCCTCA	(SEQ ID NO:732);	TAAGCCTCAA	(SEQ ID NO:733);
AAGCCTCAAT	(SEQ ID NO:734);	AGCCTCAATA	(SEQ ID NO:735);
GCCTCAATAA	(SEQ ID NO:736);	CCTCAATAAA	(SEQ ID NO:737);
CTCAATAAAG	(SEQ ID NO:738);	TCAATAAAGC	(SEQ ID NO:739);
CAATAAAGCT	(SEQ ID NO:740);	AATAAAGCTT	(SEQ ID NO:741);
ATAAAGCTTG	(SEQ ID NO:742);	TAAAGCTTGC	(SEQ ID NO:743);
AAAGCTTGCC	(SEQ ID NO:744);	AAGCTTGCCCT	(SEQ ID NO:745);
AGCTTGCCCT	(SEQ ID NO:746);	GCTTGCCCTTG	(SEQ ID NO:747);
CTTGCCCTTGA	(SEQ ID NO:748);	TTGCCCTTGAG	(SEQ ID NO:749);
TGCCTTGAGT	(SEQ ID NO:750);	GCCTTGAGTG	(SEQ ID NO:751);

- 238 -

CCTTGAGTGC	(SEQ ID NO:752);	CTTGAGTGCT	(SEQ ID NO:753);
TTGAGTGCTT	(SEQ ID NO:754);	TGAGTGCTTC	(SEQ ID NO:755);
GAGTGCTTCA	(SEQ ID NO:756);	AGTGCTTCAA	(SEQ ID NO:757);
GTGCTTCAAG	(SEQ ID NO:758);	TGCTTCAAGT	(SEQ ID NO:759);
GCTTCAAGTA	(SEQ ID NO:760);	CTTCAAGTAG	(SEQ ID NO:761);
TTCAAGTAGT	(SEQ ID NO:762);	TCAAGTAGTG	(SEQ ID NO:763);
CAAGTAGTGT	(SEQ ID NO:764);	AAGTAGTGTG	(SEQ ID NO:765);
AGTAGTGTGT	(SEQ ID NO:766);	GTAGTGTGTG	(SEQ ID NO:767);
TAGTGTGTGC	(SEQ ID NO:768);	AGTGTGTGCC	(SEQ ID NO:769);
GTGTGTGCCC	(SEQ ID NO:770);	TGTGTGCCCCG	(SEQ ID NO:771);
GTGTGCCCCGT	(SEQ ID NO:772);	TGTGCCCCGTC	(SEQ ID NO:773);
GTGCCCCGTCT	(SEQ ID NO:774);	TGCCCCGTCTG	(SEQ ID NO:775);
GCCCCGTCTGT	(SEQ ID NO:776);	CCCGTCTGTT	(SEQ ID NO:777);
CCGTCTGTTG	(SEQ ID NO:778);	CGTCTGTTGT	(SEQ ID NO:779);
GTCTGTTGTG	(SEQ ID NO:780);	TCTGTTGTGT	(SEQ ID NO:781);
CTGTTGTGTG	(SEQ ID NO:782);	TGTTGTGTGA	(SEQ ID NO:783);
GTTGTGTGAC	(SEQ ID NO:784);	TTGTGTGACT	(SEQ ID NO:785);
TGTGTGACTC	(SEQ ID NO:786);	GTGTGACTCT	(SEQ ID NO:787);
TGTGTGACTC	(SEQ ID NO:788);	GTGTGACTCT	(SEQ ID NO:789);
TGTGACTCTG	(SEQ ID NO:790);	GTGACTCTGG	(SEQ ID NO:791);
TGACTCTGGT	(SEQ ID NO:792);	GACTCTGGTA	(SEQ ID NO:793);
ACTCTGGTAA	(SEQ ID NO:794);	CTCTGGTAAC	(SEQ ID NO:795);
TCTGGTAACT	(SEQ ID NO:796);	CTGGTAACTA	(SEQ ID NO:797);
TGGTAACTAG	(SEQ ID NO:798);	GGTAACTAGA	(SEQ ID NO:799).

86. A method for determining the pathogenicity of an HIV-1 strain after said HIV-1 strain infects cells of an individual, said method comprising determining the presence of a deletion mutation in its genome of at least 10 nucleotides within the region from nucleotide 8787 to nucleotide 9709 using the nucleotide numbering of HIV-1 strain NL4-3 wherein the presence of said deletion mutation is indicative of the presence of a non-pathogenic strain of HIV-1.

87. A method according to claim 86 wherein said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

- |            |       |                |
|------------|-------|----------------|
| nucleotide | (i)   | 8830-8862;     |
|            | (ii)  | 9009-9035;     |
|            | (iii) | 9019-9029; and |
|            | (iv)  | 9033-9049.     |

88. A method according to claim 86 wherein said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

- |            |        |                |
|------------|--------|----------------|
| nucleotide | (v)    | 9281-9371;     |
|            | (vi)   | 9281-9362;     |
|            | (vii)  | 9105-9224; and |
|            | (viii) | 9271-9370.     |

89. A method according to claim 86 wherein said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

- |            |       |                |
|------------|-------|----------------|
| nucleotide | (ix)  | 8882-8928;     |
|            | (x)   | 8850-9006;     |
|            | (xi)  | 8792-9041; and |
|            | (xii) | 9112-9204.     |

90. A method according to claim 86 wherein said non-pathogenic HIV-1 carries a deletion in its genome of at least 10 nucleotides from within a region selected from the list consisting of:

- |            |        |                |
|------------|--------|----------------|
| nucleotide | (xiii) | 9105-9224;     |
|            | (xiv)  | 9389-9395; and |
|            | (xv)   | 9281-9366.     |

91. A strain of HIV-1 according to claim 86 wherein said deletion encompasses at least one of the following decanucleotides from the *nef* gene of HIV-1 strain NL4-3:

ATGGGTGGCA (SEQ ID NO:2);	TGGGTGGCAA (SEQ ID NO:3);
GGGTGGCAAG (SEQ ID NO:4);	GGTGGCAAGT (SEQ ID NO:5);

- 240 -

GTGGCAAGTG (SEQ ID NO:6);	TGGCAAGTGG (SEQ ID NO:7);
GGCAAGTGGT (SEQ ID NO:8);	GCAAGTGGTC (SEQ ID NO:9);
CAAGTGGTCA (SEQ ID NO:10);	AAGTGGTCAA (SEQ ID NO:11);
AGTGGTCAAA (SEQ ID NO:12);	GTGGTCAAAA (SEQ ID NO:13);
TGGTCAAAAA (SEQ ID NO:14);	GGTCAAAAAG (SEQ ID NO:15);
GTCAAAAAGT (SEQ ID NO:16);	TCAAAAAGTA (SEQ ID NO:17);
CAAAAAGTAG (SEQ ID NO:18);	AAAAAGTAGT (SEQ ID NO:19);
AAAAGTAGTG (SEQ ID NO:20);	AAAGTAGTGT (SEQ ID NO:21);
AAGTAGTGTG (SEQ ID NO:22);	AGTAGTGTGA (SEQ ID NO:23);
GTAGTGTGAT (SEQ ID NO:24);	TAGTGTGATT (SEQ ID NO:25);
AGTGTGATTG (SEQ ID NO:26);	GTGTGATTGG (SEQ ID NO:27);
TGTGATTGGA (SEQ ID NO:28);	GTGATTGGAT (SEQ ID NO:29);
TGATTGGATG (SEQ ID NO:30);	GATTGGATGG (SEQ ID NO:31);
ATTGGATGGC (SEQ ID NO:32);	TTGGATGGCC (SEQ ID NO:33);
TGGATGGCCT (SEQ ID NO:34);	GGATGGCCTG (SEQ ID NO:35);
GATGGCCTGC (SEQ ID NO:36);	ATGGCCTGCT (SEQ ID NO:37);
TGGCCTGCTG (SEQ ID NO:38);	GGCCTGCTGT (SEQ ID NO:39);
GCCTGCTGTA (SEQ ID NO:40);	CCTGCTGTAA (SEQ ID NO:41);
CTGCTGTAAG (SEQ ID NO:42);	TGCTGTAAGG (SEQ ID NO:43);
GCTGTAAGGG (SEQ ID NO:44);	CTGTAAGGGA (SEQ ID NO:45);
TGTAAGGGAA (SEQ ID NO:46);	GTAAGGGAAA (SEQ ID NO:47);
TAAGGGAAAG (SEQ ID NO:48);	AAGGGAAAGA (SEQ ID NO:49);
AGGGAAAGAA (SEQ ID NO:50);	GGGAAAGAAT (SEQ ID NO:51);
GGAAAGAATG (SEQ ID NO:52);	GAAAGAATGA (SEQ ID NO:53);
AAAGAATGAG (SEQ ID NO:54);	AAGAATGAGA (SEQ ID NO:55);
AGAATGAGAC (SEQ ID NO:56);	GAATGAGACG (SEQ ID NO:57);
AATGAGACGA (SEQ ID NO:58);	ATGAGACGAG (SEQ ID NO:59);
TGAGACGAGC (SEQ ID NO:60);	GAGACGAGCT (SEQ ID NO:61);
AGACGAGCTG (SEQ ID NO:62);	GACGAGCTGA (SEQ ID NO:63);
ACGAGCTGAG (SEQ ID NO:64);	CGAGCTGAGC (SEQ ID NO:65);
GAGCTGAGCC (SEQ ID NO:66);	AGCTGAGCCA (SEQ ID NO:67);
GCTGAGCCAG (SEQ ID NO:68);	CTGAGCCAGC (SEQ ID NO:69);
TGAGCCAGCA (SEQ ID NO:70);	GAGCCAGCAG (SEQ ID NO:71);
AGCCAGCAGC (SEQ ID NO:72);	GCCAGCAGCA (SEQ ID NO:73);
CCAGCAGCAG (SEQ ID NO:74);	CAGCAGCAGA (SEQ ID NO:75);
AGCAGCAGAT (SEQ ID NO:76);	GCAGCAGATG (SEQ ID NO:77);
CAGCAGATGG (SEQ ID NO:78);	AGCAGATGGG (SEQ ID NO:79);
GCAGATGGGG (SEQ ID NO:80);	CAGATGGGGT (SEQ ID NO:81);

- 241 -

AGATGGGGTG (SEQ ID NO:82);	GATGGGGTGG (SEQ ID NO:83);
ATGGGGTGGG (SEQ ID NO:84);	TGGGGTGGGA (SEQ ID NO:85);
GGGGTGGGAG (SEQ ID NO:86);	GGGTGGGAGC (SEQ ID NO:87);
GGTGGGAGCA (SEQ ID NO:88);	GTGGGAGCAG (SEQ ID NO:89);
TGGGAGCAGT (SEQ ID NO:90);	GGGAGCAGTA (SEQ ID NO:91);
GGAGCAGTAT (SEQ ID NO:92);	GAGCAGTATC (SEQ ID NO:93);
AGCAGTATCT (SEQ ID NO:94);	GCAGTATCTC (SEQ ID NO:95);
CAGTATCTCG (SEQ ID NO:96);	AGTATCTCGA (SEQ ID NO:97);
GTATCTCGAG (SEQ ID NO:98);	TATCTCGAGA (SEQ ID NO:99);
ATCTCGAGAC (SEQ ID NO:100);	TCTCGAGACC (SEQ ID NO:101);
CTCGAGACCT (SEQ ID NO:102);	TCGAGACCTA (SEQ ID NO:103);
CGAGACCTAG (SEQ ID NO:104);	GAGACCTAGA (SEQ ID NO:105);
AGACCTAGAA (SEQ ID NO:106);	GACCTAGAAA (SEQ ID NO:107);
ACCTAGAAAA (SEQ ID NO:108);	CCTAGAAAAA (SEQ ID NO:109);
CTAGAAAAAC (SEQ ID NO:110);	TAGAAAAACA (SEQ ID NO:111);
AGAAAAACAT (SEQ ID NO:112);	GAAAAACATG (SEQ ID NO:113);
AAAAACATGG (SEQ ID NO:114);	AAACATGGA (SEQ ID NO:115);
AAACATGGAG (SEQ ID NO:116);	AACATGGAGC (SEQ ID NO:117);
ACATGGAGCA (SEQ ID NO:118);	CATGGAGCAA (SEQ ID NO:119);
ATGGAGCAAT (SEQ ID NO:120);	TGGAGCAATC (SEQ ID NO:121);
GGAGCAATCA (SEQ ID NO:122);	GAGCAATCAC (SEQ ID NO:123);
AGCAATCACA (SEQ ID NO:124);	GCAATCACAA (SEQ ID NO:125);
CAATCACAAG (SEQ ID NO:126);	AATCACAAGT (SEQ ID NO:127);
ATCACAAGTA (SEQ ID NO:128);	TCACAAGTAG (SEQ ID NO:129);
CACAAGTAGC (SEQ ID NO:130);	ACAAGTAGCA (SEQ ID NO:131);
CAAGTAGCAA (SEQ ID NO:132);	AAGTAGCAAT (SEQ ID NO:133);
AGTAGCAATA (SEQ ID NO:134);	GTAGCAATAC (SEQ ID NO:135);
TAGCAATACA (SEQ ID NO:136);	AGCAATACAG (SEQ ID NO:137);
GCAATACAGC (SEQ ID NO:138);	CAATACAGCA (SEQ ID NO:139);
AATACAGCAG (SEQ ID NO:140);	ATACAGCAGC (SEQ ID NO:141);
TACAGCAGCT (SEQ ID NO:142);	ACAGCAGCTA (SEQ ID NO:143);
CAGCAGCTAA (SEQ ID NO:144);	AGCAGCTAAC (SEQ ID NO:145);
GCAGCTAACA (SEQ ID NO:146);	CAGCTAACAA (SEQ ID NO:147);
AGCTAACAAT (SEQ ID NO:148);	GCTAACAATG (SEQ ID NO:149);
CTAACAATGC (SEQ ID NO:150);	TAACAATGCT (SEQ ID NO:151);
AACAATGCTG (SEQ ID NO:152);	ACAATGCTGC (SEQ ID NO:153);
CAATGCTGCT (SEQ ID NO:154);	AATGCTGCTT (SEQ ID NO:155);
ATGCTGCTTG (SEQ ID NO:156);	TGCTGCTTGT (SEQ ID NO:157);

- 242 -

GCTGCTTGTG (SEQ ID NO:158);	CTGCTTGTGC (SEQ ID NO:159);
TGCTTGTGCC (SEQ ID NO:160);	GCTTGTGCCT (SEQ ID NO:161);
CTTGTGCCTG (SEQ ID NO:162);	TTGTGCCTGG (SEQ ID NO:163);
TGTGCCTGGC (SEQ ID NO:164);	GTGCCTGGCT (SEQ ID NO:165);
TGCCTGGCTA (SEQ ID NO:166);	GCCTGGCTAG (SEQ ID NO:167);
CCTGGCTAGA (SEQ ID NO:168);	CTGGCTAGAA (SEQ ID NO:169);
TGGCTAGAAG (SEQ ID NO:170);	GGCTAGAAGC (SEQ ID NO:171);
GCTAGAAGCA (SEQ ID NO:172);	CTAGAAGCAC (SEQ ID NO:173);
TAGAAGCACA (SEQ ID NO:174);	AGAAGCACAA (SEQ ID NO:175);
GAAGCACAAG (SEQ ID NO:176);	AAGCACAAGA (SEQ ID NO:177);
AGCACAAGAG (SEQ ID NO:178);	GCACAAGAGG (SEQ ID NO:179);
CACAAGAGGA (SEQ ID NO:180);	ACAAGAGGAG (SEQ ID NO:181);
CAAGAGGAGG (SEQ ID NO:182);	AAGAGGAGGA (SEQ ID NO:183);
AGAGGAGGAA (SEQ ID NO:184);	GAGGAGGAAG (SEQ ID NO:185);
AGGAGGAAGA (SEQ ID NO:186);	GGAGGAAGAG (SEQ ID NO:187);
GAGGAAGAGG (SEQ ID NO:188);	AGGAAGAGGT (SEQ ID NO:189);
GGAAGAGGTG (SEQ ID NO:190);	GAAGAGGTGG (SEQ ID NO:191);
AAGAGGTGGG (SEQ ID NO:192);	AGAGGTGGGT (SEQ ID NO:193);
GAGGTGGGTT (SEQ ID NO:194);	AGGTGGGTTT (SEQ ID NO:195);
GGTGGGTTTT (SEQ ID NO:196);	GTGGGTTTTC (SEQ ID NO:197);
TGGGTTTTCC (SEQ ID NO:198);	GGGTTTTCCA (SEQ ID NO:199);
GGTTTTCCAG (SEQ ID NO:200);	GTTTTCCAGT (SEQ ID NO:201);
TTTTCCAGTC (SEQ ID NO:202);	TTTCCAGTCA (SEQ ID NO:203);
TTCCAGTCAC (SEQ ID NO:204);	TCCAGTCACA (SEQ ID NO:205);
CCAGTCACAC (SEQ ID NO:206);	CAGTCACACC (SEQ ID NO:207);
AGTCACACCT (SEQ ID NO:208);	GTCACACCTC (SEQ ID NO:209);
TCACACCTCA (SEQ ID NO:210);	CACACCTCAG (SEQ ID NO:211);
ACACCTCAGG (SEQ ID NO:212);	CACCTCAGGT (SEQ ID NO:213);
ACCTCAGGTA (SEQ ID NO:214);	CCTCAGGTAC (SEQ ID NO:215);
CTCAGGTACC (SEQ ID NO:216);	TCAGGTACCT (SEQ ID NO:217);
CAGGTACCTT (SEQ ID NO:218);	AGGTACCTTT (SEQ ID NO:219);
GGTACCTTTA (SEQ ID NO:220);	GTACCTTTAA (SEQ ID NO:221);
TACCTTTAAG (SEQ ID NO:222);	ACCTTTAAGA (SEQ ID NO:223);
CCTTTAAGAC (SEQ ID NO:224);	CTTTAAGACC (SEQ ID NO:225);
TTTAAGACCA (SEQ ID NO:226);	TTAAGACCAA (SEQ ID NO:227);
TAAGACCAAT (SEQ ID NO:228);	AAGACCAATG (SEQ ID NO:229);
AGACCAATGA (SEQ ID NO:230);	GACCAATGAC (SEQ ID NO:231);
ACCAATGACT (SEQ ID NO:232);	CCAATGACTT (SEQ ID NO:233);



- 243 -

CAATGACTTA (SEQ ID NO:234);	AATGACTTAC (SEQ ID NO:235);
ATGACTTACA (SEQ ID NO:236);	TGACTTACAA (SEQ ID NO:237);
GACTTACAAG (SEQ ID NO:238);	ACTTACAAGG (SEQ ID NO:239);
CTTACAAGGC (SEQ ID NO:240);	TTACAAGGCA (SEQ ID NO:241);
TACAAGGCAG (SEQ ID NO:242);	ACAAGGCAGC (SEQ ID NO:243);
CAAGGCAGCT (SEQ ID NO:244);	AAGGCAGCTG (SEQ ID NO:245);
AGGCAGCTGT (SEQ ID NO:246);	GGCAGCTGTA (SEQ ID NO:247);
GCAGCTGTAG (SEQ ID NO:248);	CAGCTGTAGA (SEQ ID NO:249);
AGCTGTAGAT (SEQ ID NO:250);	GCTGTAGATC (SEQ ID NO:251);
CTGTAGATCT (SEQ ID NO:252);	TGTAGATCTT (SEQ ID NO:253);
GTAGATCTTA (SEQ ID NO:254);	TAGATCTTAG (SEQ ID NO:255);
AGATCTTAGC (SEQ ID NO:256);	GATCTTAGCC (SEQ ID NO:257);
ATCTTAGCCA (SEQ ID NO:258);	TCTTAGCCAC (SEQ ID NO:259);
CTTAGCCACT (SEQ ID NO:260);	TTAGCCACTT (SEQ ID NO:261);
TAGCCACTTT (SEQ ID NO:262);	AGCCACTTTT (SEQ ID NO:263);
GCCACTTTTT (SEQ ID NO:264);	CCACTTTTTA (SEQ ID NO:265);
CACTTTTTTAA (SEQ ID NO:266);	ACTTTTTTAAA (SEQ ID NO:267);
CTTTTTTAAAA (SEQ ID NO:268);	TTTTTTAAAAG (SEQ ID NO:269);
TTTTTAAAAGA (SEQ ID NO:270);	TTTTTAAAAGAA (SEQ ID NO:271);
TTAAAAGAAA (SEQ ID NO:272);	TAAAAGAAAA (SEQ ID NO:273);
AAAAGAAAAG (SEQ ID NO:274);	AAAGAAAAGG (SEQ ID NO:275);
AAGAAAAGGG (SEQ ID NO:276);	AGAAAAGGGG (SEQ ID NO:277);
GAAAAGGGGG (SEQ ID NO:278);	AAAAGGGGGG (SEQ ID NO:279);
AAAGGGGGGA (SEQ ID NO:280);	AAGGGGGGAC (SEQ ID NO:281);
AGGGGGGACT (SEQ ID NO:282);	GGGGGGGACTG (SEQ ID NO:283);
GGGGGACTGG (SEQ ID NO:284);	GGGGACTGGA (SEQ ID NO:285);
GGGACTGGA (SEQ ID NO:286);	GGACTGGAAG (SEQ ID NO:287);
GACTGGAAGG (SEQ ID NO:288);	ACTGGAAGGG (SEQ ID NO:289);
CTGGAAGGGC (SEQ ID NO:290);	TGGAAGGGCT (SEQ ID NO:291);
GGAAGGGCTA (SEQ ID NO:292);	GAAGGGCTAA (SEQ ID NO:293);
AAGGGCTAAT (SEQ ID NO:294);	AGGGCTAATT (SEQ ID NO:295);
GGGCTAATTC (SEQ ID NO:296);	GGCTAATTCA (SEQ ID NO:297);
GCTAATTCAC (SEQ ID NO:298);	CTAATTCACT (SEQ ID NO:299);
TAATTCACTC (SEQ ID NO:300);	AATTCACTCC (SEQ ID NO:301);
ATTCACTCCC (SEQ ID NO:302);	TTCACTCCCA (SEQ ID NO:303);
TCACTCCCAA (SEQ ID NO:304);	CACTCCCAAA (SEQ ID NO:305);
ACTCCCAAAG (SEQ ID NO:306);	CTCCCAAAGA (SEQ ID NO:307);
TCCCAAAGAA (SEQ ID NO:308);	CCCAAAGAAG (SEQ ID NO:309);

CCAAAGAAGA (SEQ ID NO:310);	CAAAGAAGAC (SEQ ID NO:311);
AAAGAAGACA (SEQ ID NO:312);	AAGAAGACAA (SEQ ID NO:313);
AGAAGACAAG (SEQ ID NO:314);	GAAGACAAGA (SEQ ID NO:315);
AAGACAAGAT (SEQ ID NO:316);	AGACAAGATA (SEQ ID NO:317);
GACAAGATAT (SEQ ID NO:318);	ACAAGATATC (SEQ ID NO:319);
CAAGATATCC (SEQ ID NO:320);	AAGATATCCT (SEQ ID NO:321);
AGATATCCTT (SEQ ID NO:322);	GATATCCTTG (SEQ ID NO:323);
ATATCCTTGA (SEQ ID NO:324);	TATCCTTGAT (SEQ ID NO:325);
ATCCTTGATC (SEQ ID NO:326);	TCCTTGATCT (SEQ ID NO:327);
CCTTGATCTG (SEQ ID NO:328);	CTTGATCTGT (SEQ ID NO:329);
TTGATCTGTG (SEQ ID NO:330);	TGATCTGTGG (SEQ ID NO:331);
GATCTGTGGA (SEQ ID NO:332);	ATCTGTGGAT (SEQ ID NO:333);
TCTGTGGATC (SEQ ID NO:334);	CTGTGGATCT (SEQ ID NO:335);
TGTGGATCTA (SEQ ID NO:336);	GTGGATCTAC (SEQ ID NO:337);
TGGATCTACC (SEQ ID NO:338);	GGATCTACCA (SEQ ID NO:339);
GATCTACCAC (SEQ ID NO:340);	ATCTACCACA (SEQ ID NO:341);
TCTACCACAC (SEQ ID NO:342);	CTACCACACA (SEQ ID NO:343);
TACCACACAC (SEQ ID NO:344);	ACCACACACA (SEQ ID NO:345);
CCACACACAA (SEQ ID NO:346);	CACACACAAG (SEQ ID NO:347);
ACACACAAGG (SEQ ID NO:348);	CACACAAGGC (SEQ ID NO:349);
ACACAAGGCT (SEQ ID NO:350);	CACAAGGCTA (SEQ ID NO:351);
ACAAGGCTAC (SEQ ID NO:352);	CAAGGCTACT (SEQ ID NO:353);
AAGGCTACTT (SEQ ID NO:354);	AGGCTACTTC (SEQ ID NO:355);
GGCTACTTCC (SEQ ID NO:356);	GCTACTTCCC (SEQ ID NO:357);
CTACTTCCCT (SEQ ID NO:358);	TACTTCCCTG (SEQ ID NO:359);
ACTTCCCTGA (SEQ ID NO:360);	CTTCCCTGAT (SEQ ID NO:361);
TTCCCTGATT (SEQ ID NO:362);	TCCCTGATTG (SEQ ID NO:363);
CCCTGATTGG (SEQ ID NO:364);	CCTGATTGGC (SEQ ID NO:365);
CTGATTGGCA (SEQ ID NO:366);	TGATTGGCAG (SEQ ID NO:367);
GATTGGCAGA (SEQ ID NO:368);	ATTGGCAGAA (SEQ ID NO:369);
TTGGCAGAAC (SEQ ID NO:370);	TGGCAGAACT (SEQ ID NO:371);
GGCAGAACTA (SEQ ID NO:372);	GCAGAACTAC (SEQ ID NO:373);
CAGAACTACA (SEQ ID NO:374);	AGAACTACAC (SEQ ID NO:375);
GAACCTACACA (SEQ ID NO:376);	AACTACACAC (SEQ ID NO:377);
ACTACACACC (SEQ ID NO:378);	CTACACACCA (SEQ ID NO:379);
TACACACCAG (SEQ ID NO:380);	ACACACCAGG (SEQ ID NO:381);
CACACCAGGG (SEQ ID NO:382);	ACACCAGGGC (SEQ ID NO:383);
CACCAGGGCC (SEQ ID NO:384);	ACCAGGGCCA (SEQ ID NO:385);

- 245 -

CCAGGGCCAG (SEQ ID NO:386);	CAGGGCCAGG (SEQ ID NO:387);
AGGGCCAGGG (SEQ ID NO:388);	GGGCCAGGGG (SEQ ID NO:389);
GGCCAGGGGT (SEQ ID NO:390);	GCCAGGGGTC (SEQ ID NO:391);
CCAGGGGTCA (SEQ ID NO:392);	CAGGGGTCA (SEQ ID NO:393);
AGGGGTCA (SEQ ID NO:394);	GGGGTCA (SEQ ID NO:395);
GGGTCA (SEQ ID NO:396);	GGTCA (SEQ ID NO:397);
GTCAGATATC (SEQ ID NO:398);	TCAGATATCC (SEQ ID NO:399);
CAGATATCCA (SEQ ID NO:400);	AGATATCCAC (SEQ ID NO:401);
GATATCCACT (SEQ ID NO:402);	ATATCCACTG (SEQ ID NO:403);
TATCCACTGA (SEQ ID NO:404);	ATCCACTGAC (SEQ ID NO:405);
TCCACTGACC (SEQ ID NO:406);	CCACTGACCT (SEQ ID NO:407);
CACTGACCTT (SEQ ID NO:408);	ACTGACCTTT (SEQ ID NO:409);
CTGACCTTTG (SEQ ID NO:410);	TGACCTTTGG (SEQ ID NO:411);
GACCTTTGGA (SEQ ID NO:412);	ACCTTTGGAT (SEQ ID NO:413);
CCTTTGGATG (SEQ ID NO:414);	CTTTGGATGG (SEQ ID NO:415);
TTTGGATGGT (SEQ ID NO:416);	TTGGATGGTG (SEQ ID NO:417);
TGGATGGTGC (SEQ ID NO:418);	GGATGGTGCT (SEQ ID NO:419);
GATGGTGCTA (SEQ ID NO:420);	ATGGTGCTAC (SEQ ID NO:421);
TGGTGCTACA (SEQ ID NO:422);	GGTGCTACAA (SEQ ID NO:423);
GTGCTACAAG (SEQ ID NO:424);	TGCTACAAGC (SEQ ID NO:425);
GCTACAAGCT (SEQ ID NO:426);	CTACAAGCTA (SEQ ID NO:427);
TACAAGCTAG (SEQ ID NO:428);	ACAAGCTAGT (SEQ ID NO:429);
CAAGCTAGTA (SEQ ID NO:430);	AAGCTAGTAC (SEQ ID NO:431);
AGCTAGTACC (SEQ ID NO:432);	GCTAGTACCA (SEQ ID NO:433);
CTAGTACCAG (SEQ ID NO:434);	TAGTACCAGT (SEQ ID NO:435);
AGTACCAGTT (SEQ ID NO:436);	GTACCAGTTG (SEQ ID NO:437);
TACCAGTTGA (SEQ ID NO:438);	ACCAGTTGAG (SEQ ID NO:439);
CCAGTTGAGC (SEQ ID NO:440);	CAGTTGAGCC (SEQ ID NO:441);
AGTTGAGCCA (SEQ ID NO:442);	GTTGAGCCAG (SEQ ID NO:443);
TTGAGCCAGA (SEQ ID NO:444);	TGAGCCAGAT (SEQ ID NO:445);
GAGCCAGATA (SEQ ID NO:446);	AGCCAGATAA (SEQ ID NO:447);
GCCAGATAAG (SEQ ID NO:448);	CCAGATAAGG (SEQ ID NO:449);
CAGATAAGGT (SEQ ID NO:450);	AGATAAGGTA (SEQ ID NO:451);
GATAAGGTAG (SEQ ID NO:452);	ATAAGGTAGA (SEQ ID NO:453);
TAAGGTAGAA (SEQ ID NO:454);	AAGGTAGAA (SEQ ID NO:455);
AGGTAGAA (SEQ ID NO:456);	GGTAGAA (SEQ ID NO:457);
GTAGAAGAGG (SEQ ID NO:458);	TAGAAGAGGC (SEQ ID NO:459);
AGAAGAGGCC (SEQ ID NO:460);	GAAGAGGCCA (SEQ ID NO:461);

- 246 -

AAGAGGCCAA (SEQ ID NO:462);	AGAGGCCAAT (SEQ ID NO:463);
GAGGCCAATA (SEQ ID NO:464);	AGGCCAATAA (SEQ ID NO:465);
GGCCAATAAA (SEQ ID NO:466);	GCCAATAAAG (SEQ ID NO:467);
CCAATAAAGG (SEQ ID NO:468);	CAATAAAGGA (SEQ ID NO:469);
AATAAAGGAG (SEQ ID NO:470);	ATAAAGGAGA (SEQ ID NO:471);
TAAAGGAGAG (SEQ ID NO:472);	AAAGGAGAGA (SEQ ID NO:473);
AAGGAGAGAA (SEQ ID NO:474);	AGGAGAGAAC (SEQ ID NO:475);
GGAGAGAACA (SEQ ID NO:476);	GAGAGAACAC (SEQ ID NO:477);
AGAGAACACC (SEQ ID NO:478);	GAGAACACCA (SEQ ID NO:479);
AGAACACCAG (SEQ ID NO:480);	GAACACCAGC (SEQ ID NO:481);
AACACCAGCT (SEQ ID NO:482);	ACACCAGCTT (SEQ ID NO:483);
CACCAGCTTG (SEQ ID NO:484);	ACCAGCTTGT (SEQ ID NO:485);
CCAGCTTGTT (SEQ ID NO:486);	CAGCTTGTTA (SEQ ID NO:487);
AGCTTGTTAC (SEQ ID NO:488);	GCTTGTTACA (SEQ ID NO:489);
CTTGTTACAC (SEQ ID NO:490);	TTGTTACACC (SEQ ID NO:491);
TGTTACACCC (SEQ ID NO:492);	GTTACACCCT (SEQ ID NO:493);
TTACACCCTG (SEQ ID NO:494);	TACACCCTGT (SEQ ID NO:495);
ACACCCTGTG (SEQ ID NO:496);	CACCCTGTGA (SEQ ID NO:497);
ACCCTGTGAG (SEQ ID NO:498);	CCCTGTGAGC (SEQ ID NO:499);
CCTGTGAGCC (SEQ ID NO:500);	CTGTGAGCCT (SEQ ID NO:501);
TGTGAGCCTG (SEQ ID NO:502);	GTGAGCCTGC (SEQ ID NO:503);
TGAGCCTGCA (SEQ ID NO:504);	GAGCCTGCAT (SEQ ID NO:505);
AGCCTGCATG (SEQ ID NO:506);	GCCTGCATGG (SEQ ID NO:507);
CCTGCATGGA (SEQ ID NO:508);	CTGCATGGAA (SEQ ID NO:509);
TGCATGGAAT (SEQ ID NO:510);	GCATGGAATG (SEQ ID NO:511);
CATGGAATGG (SEQ ID NO:512);	ATGGAATGGA (SEQ ID NO:513);
TGGAATGGAT (SEQ ID NO:514);	GGAATGGATG (SEQ ID NO:515);
GAATGGATGA (SEQ ID NO:516);	AATGGATGAC (SEQ ID NO:517);
ATGGATGACC (SEQ ID NO:518);	TGGATGACCC (SEQ ID NO:519);
GGATGACCCT (SEQ ID NO:520);	GATGACCCTG (SEQ ID NO:521);
ATGACCCTGA (SEQ ID NO:522);	TGACCCTGAG (SEQ ID NO:523);
GACCCTGAGA (SEQ ID NO:524);	ACCCTGAGAG (SEQ ID NO:525);
CCCTGAGAGA (SEQ ID NO:526);	CCTGAGAGAG (SEQ ID NO:527);
CTGAGAGAGA (SEQ ID NO:528);	TGAGAGAGAA (SEQ ID NO:529);
GAGAGAGAAG (SEQ ID NO:530);	AGAGAGAAGT (SEQ ID NO:531);
GAGAGAAGTG (SEQ ID NO:532);	AGAGAAGTGT (SEQ ID NO:533);
GAGAAGTGTT (SEQ ID NO:534);	AGAAGTGTTA (SEQ ID NO:535);
GAAGTGTTAG (SEQ ID NO:536);	AAGTGTTAGA (SEQ ID NO:537);

AGTGTTAGAG (SEQ ID NO:538);	GTGTTAGAGT (SEQ ID NO:539);
TGTTAGAGTG (SEQ ID NO:540);	GTTAGAGTGG (SEQ ID NO:541);
TTAGAGTGGA (SEQ ID NO:542);	TAGAGTGGAG (SEQ ID NO:543);
AGAGTGGAGG (SEQ ID NO:544);	GAGTGGAGGT (SEQ ID NO:545);
AGTGGAGGTT (SEQ ID NO:546);	GTGGAGGTTT (SEQ ID NO:547);
TGGAGGTTTG (SEQ ID NO:548);	GGAGGTTTGA (SEQ ID NO:549);
GAGGTTTGAC (SEQ ID NO:550);	AGGTTTGACA (SEQ ID NO:551);
GGTTTGACAG (SEQ ID NO:552);	GTTTGACAGC (SEQ ID NO:553);
TTTGACAGCC (SEQ ID NO:554);	TTGACAGCCG (SEQ ID NO:555);
TGACAGCCGC (SEQ ID NO:556);	GACAGCCGCC (SEQ ID NO:557);
ACAGCCGCCT (SEQ ID NO:558);	CAGCCGCCTA (SEQ ID NO:559);
AGCCGCCTAG (SEQ ID NO:560);	GCCGCCTAGC (SEQ ID NO:561);
CCGCCTAGCA (SEQ ID NO:562);	CGCCTAGCAT (SEQ ID NO:563);
GCCTAGCATT (SEQ ID NO:564);	CCTAGCATTT (SEQ ID NO:565);
CTAGCATTTT (SEQ ID NO:566);	TAGCATTTCA (SEQ ID NO:567);
AGCATTTTCA (SEQ ID NO:568);	GCATTTTCATC (SEQ ID NO:569);
CATTTTCATCA (SEQ ID NO:570);	ATTTTCATCAC (SEQ ID NO:571);
TTTCATCACG (SEQ ID NO:572);	TTCATCACGT (SEQ ID NO:573);
TCATCACGTG (SEQ ID NO:574);	CATCACGTGG (SEQ ID NO:575);
ATCACGTGGC (SEQ ID NO:576);	TCACGTGGCC (SEQ ID NO:577);
CACGTGGCCC (SEQ ID NO:578);	ACGTGGCCCC (SEQ ID NO:579);
CGTGGCCCCG (SEQ ID NO:580);	GTGGCCCCGAG (SEQ ID NO:581);
TGGCCCCGAG (SEQ ID NO:582);	GGCCCCGAGAG (SEQ ID NO:583);
GCCCCGAGAGC (SEQ ID NO:584);	CCCGAGAGCT (SEQ ID NO:585);
CCGAGAGCTG (SEQ ID NO:586);	CGAGAGCTGC (SEQ ID NO:587);
GAGAGCTGCA (SEQ ID NO:588);	AGAGCTGCAT (SEQ ID NO:589);
GAGCTGCATC (SEQ ID NO:590);	AGCTGCATCC (SEQ ID NO:591);
GCTGCATCCG (SEQ ID NO:592);	CTGCATCCGG (SEQ ID NO:593);
TGCATCCGGA (SEQ ID NO:594);	GCATCCGGAG (SEQ ID NO:595);
CATCCGGAGT (SEQ ID NO:596);	ATCCGGAGTA (SEQ ID NO:597);
TCCGGAGTAC (SEQ ID NO:598);	CCGGAGTACT (SEQ ID NO:599);
CGGAGTACTT (SEQ ID NO:600);	GGAGTACTTC (SEQ ID NO:601);
GAGTACTTCA (SEQ ID NO:602);	AGTACTTCAA (SEQ ID NO:603);
GTA TTTCAAG (SEQ ID NO:604);	TACTTCAAGA (SEQ ID NO:605);
ACTTCAAGAA (SEQ ID NO:606);	CTTCAAGAAC (SEQ ID NO:607);
TTCAAGAACT (SEQ ID NO:608);	TCAAGAACTG (SEQ ID NO:609);
CAAGAACTGC (SEQ ID NO:610);	AAGAACTGCT (SEQ ID NO:611);
AGAACTGCTG (SEQ ID NO:612);	GAACTGCTGA (SEQ ID NO:613);

92. A method according to claim 86 wherein said deletion encompasses at least one of the following decanucleotides from the LTR region of HIV-1 strain NL4-3:

GCTTTTGGCC	(SEQ ID NO:652);	CTTTTGCCT	(SEQ ID NO:653);
TTTTTGCCTG	(SEQ ID NO:654);	TTTTGCCTGT	(SEQ ID NO:655);
TTTGCCTGTA	(SEQ ID NO:656);	TTGCCTGTAC	(SEQ ID NO:657);
TGCCTGTACT	(SEQ ID NO:658);	GCCTGTACTG	(SEQ ID NO:659);
CCTGTACTGG	(SEQ ID NO:660);	CTGTACTGGG	(SEQ ID NO:661);
TGTACTGGGT	(SEQ ID NO:662);	GTACTGGGTC	(SEQ ID NO:663);
TACTGGGTCT	(SEQ ID NO:664);	ACTGGGTCTC	(SEQ ID NO:665);
CTGGGTCTCT	(SEQ ID NO:666);	TGGGTCTCTC	(SEQ ID NO:667);
GGGTCTCTCT	(SEQ ID NO:668);	GGTCTCTCTG	(SEQ ID NO:669);
GTCTCTCTGG	(SEQ ID NO:670);	TCTCTCTGGT	(SEQ ID NO:671);
CTCTCTGGTT	(SEQ ID NO:672);	TCTCTGGTTA	(SEQ ID NO:673);
CTCTGGTTAG	(SEQ ID NO:674);	TCTCTGGTTA	(SEQ ID NO:675);
CTGGTTAGAC	(SEQ ID NO:676);	TGGTTAGACC	(SEQ ID NO:677);
GGTTAGACCA	(SEQ ID NO:678);	GTTAGACCAG	(SEQ ID NO:679);
TTAGACCAGA	(SEQ ID NO:680);	TAGACCAGAT	(SEQ ID NO:681);
AGACCAGATC	(SEQ ID NO:682);	GACCAGATCT	(SEQ ID NO:683);
ACCAGATCTG	(SEQ ID NO:684);	CCAGATCTGA	(SEQ ID NO:685);
CAGATCTGAG	(SEQ ID NO:686);	AGATCTGAGC	(SEQ ID NO:687);
GATCTGAGCC	(SEQ ID NO:688);	ATCTGAGCCT	(SEQ ID NO:689);
TCTGAGCCTG	(SEQ ID NO:690);	CTGAGCCTGG	(SEQ ID NO:691);
TGAGCCTGGG	(SEQ ID NO:692);	GAGCCTGGGA	(SEQ ID NO:693);
AGCCTGGGAG	(SEQ ID NO:694);	GCCTGGGAGC	(SEQ ID NO:695);
CCTGGGAGCT	(SEQ ID NO:696);	CTGGGAGCTC	(SEQ ID NO:697);
TGGGAGCTCT	(SEQ ID NO:698);	GGGAGCTCTC	(SEQ ID NO:699);
GGAGCTCTCT	(SEQ ID NO:700);	GAGCTCTCTG	(SEQ ID NO:701);
AGCTCTCTGG	(SEQ ID NO:702);	GCTCTCTGGC	(SEQ ID NO:703);
CTCTCTGGCT	(SEQ ID NO:704);	TCTCTGGCTA	(SEQ ID NO:705);
CTCTGGCTAA	(SEQ ID NO:706);	TCTGGCTAAC	(SEQ ID NO:707);
CTGGCTAACT	(SEQ ID NO:708);	TGGCTAACTA	(SEQ ID NO:709);
GGCTAACTAG	(SEQ ID NO:710);	GCTAACTAGG	(SEQ ID NO:711);
CTAACTAGGG	(SEQ ID NO:712);	TAACTAGGGA	(SEQ ID NO:713);
AACTAGGGAA	(SEQ ID NO:714);	ACTAGGGAAC	(SEQ ID NO:715);
CTAGGGAACC	(SEQ ID NO:716);	TAGGGAACCC	(SEQ ID NO:717);
AGGGAACCCA	(SEQ ID NO:718);	GGAACCCAC	(SEQ ID NO:719);

GGAACCCACT	(SEQ ID NO:720);	GAACCCACTG	(SEQ ID NO:721);
AACCCACTGC	(SEQ ID NO:722);	ACCCACTGCT	(SEQ ID NO:723);
CCCACTGCTT	(SEQ ID NO:724);	CCACTGCTTA	(SEQ ID NO:725);
CACTGCTTAA	(SEQ ID NO:726);	ACTGCTTAAG	(SEQ ID NO:727);
CTGCTTAAGC	(SEQ ID NO:728);	TGCTTAAGCC	(SEQ ID NO:729);
GCTTAAGCCT	(SEQ ID NO:730);	CTTAAGCCTC	(SEQ ID NO:731);
TTAAGCCTCA	(SEQ ID NO:732);	TAAGCCTCAA	(SEQ ID NO:733);
AAGCCTCAAT	(SEQ ID NO:734);	AGCCTCAATA	(SEQ ID NO:735);
GCCTCAATAA	(SEQ ID NO:736);	CCTCAATAAA	(SEQ ID NO:737);
CTCAATAAAG	(SEQ ID NO:738);	TCAATAAAGC	(SEQ ID NO:739);
CAATAAAGCT	(SEQ ID NO:740);	AATAAAGCTT	(SEQ ID NO:741);
ATAAAGCTTG	(SEQ ID NO:742);	TAAAGCTTGC	(SEQ ID NO:743);
AAAGCTTGCC	(SEQ ID NO:744);	AAGCTTGCCCT	(SEQ ID NO:745);
AGCTTGCCCT	(SEQ ID NO:746);	GCTTGCCCTTG	(SEQ ID NO:747);
CTTGCCCTGA	(SEQ ID NO:748);	TTGCCCTTGAG	(SEQ ID NO:749);
TGCCCTTGAGT	(SEQ ID NO:750);	GCCTTGAGTG	(SEQ ID NO:751);
CCTTGAGTGC	(SEQ ID NO:752);	CTTGAGTGCT	(SEQ ID NO:753);
TTGAGTGCTT	(SEQ ID NO:754);	TGAGTGCTTC	(SEQ ID NO:755);
GAGTGCTTCA	(SEQ ID NO:756);	AGTGCTTCAA	(SEQ ID NO:757);
GTGCTTCAAG	(SEQ ID NO:758);	TGCTTCAAGT	(SEQ ID NO:759);
GCTTCAAGTA	(SEQ ID NO:760);	CTTCAAGTAG	(SEQ ID NO:761);
TTCAAGTAGT	(SEQ ID NO:762);	TCAAGTAGTG	(SEQ ID NO:763);
CAAGTAGTGT	(SEQ ID NO:764);	AAGTAGTGTG	(SEQ ID NO:765);
AGTAGTGTGT	(SEQ ID NO:766);	GTAGTGTGTG	(SEQ ID NO:767);
TAGTGTGTGC	(SEQ ID NO:768);	AGTGTGTGCC	(SEQ ID NO:769);
GTGTGTGCCC	(SEQ ID NO:770);	TGTGTGCCCCG	(SEQ ID NO:771);
GTGTGCCCCGT	(SEQ ID NO:772);	TGTGCCCCGTC	(SEQ ID NO:773);
GTGCCCCGTCT	(SEQ ID NO:774);	TGCCCCGTCTG	(SEQ ID NO:775);
GCCCGTCTGT	(SEQ ID NO:776);	CCCGTCTGTT	(SEQ ID NO:777);
CCGTCTGTTG	(SEQ ID NO:778);	CGTCTGTTGT	(SEQ ID NO:779);
GTCTGTTGTG	(SEQ ID NO:780);	TCTGTTGTGT	(SEQ ID NO:781);
CTGTTGTGTG	(SEQ ID NO:782);	TGTTGTGTGA	(SEQ ID NO:783);
GTTGTGTGAC	(SEQ ID NO:784);	TTGTGTGACT	(SEQ ID NO:785);
TGTGTGACTC	(SEQ ID NO:786);	GTGTGACTCT	(SEQ ID NO:787);
TGTGTGACTC	(SEQ ID NO:788);	GTGTGACTCT	(SEQ ID NO:789);
TGTGACTCTG	(SEQ ID NO:790);	GTGACTCTGG	(SEQ ID NO:791);

TGACTCTGGT (SEQ ID NO:792); GACTCTGGTA (SEQ ID NO:793);  
ACTCTGGTAA (SEQ ID NO:794); CTCTGGTAAC (SEQ ID NO:795);  
TCTGGTAACT (SEQ ID NO:796); CTGGTAACTA (SEQ ID NO:797);  
TGGTAACTAG (SEQ ID NO:798); GGTAAC TAGA (SEQ ID NO:799).



FIGURE 1

1/47

2/47
3/47
4/47
5/47
6/47
7/47
8/47
9/47
10/47
11/47
12/47
13/47
14/47
15/47
16/47
17/47
18/47
19/47

2/47

## FIGURE 1

8072

NL43 GAACAGATTGGAAATAACATGACCTGGATGGAGTGGACAGAGAAATTAA 8121  
\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*  
D36P GAAGAGATTGGGAGAACATGACCTGGATGCAGTGGGAAAAGAAATTCA  
\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*  
C18S GAAACAAATTGGGATAACATGACCTGGATGCAGTGGGAAAGAGAAATTGA  
\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*  
C18M GAAACAAATTGGGATAACATGACCTGGATGCAGTGGGAAAGAGAAATTGA  
\*\*\*\*\*  
C98H GAAATTAA

8171

NL43 CAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGC  
\*\*\* \*\*\*\*\* \* \*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*  
D36P CAATCACACAAAATACATATACTCCTTACTTGAAATAATCGCAGAACCAAC  
\*\*\*\*\* \* \*\*\*\*\* \* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*  
C18S CAATTACACAACATAATATACACCTTAATTGAAGAATCGCAGAACCAAC  
\*\*\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*  
C18M CAATTACACAACATAATATACACCTTAATTGAAGAATCGCAGAACCAAC  
\*\*\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \* \*\*\*\*\* \*  
C98H CAATTACACAAGATTAAATATACAACCTTAATTGAAGAATCGCAGAACCAAC

3/47

## FIGURE 1

NL43 AAGAAAAGAAATGAACAAGAAATTATTGGAATTAGATAAATGGGCAAGTTTG 8221  
\*\*\*\*\*  
D36P AAGAAAAGAAATGAACAAGAACTATTGGAATTGGATCAATGGGCAAGTTTG  
\*\*\*\*\*  
C18S AAGAAAAGAAATGAAC TAGAAATTATTGGAATTGGATAAATGGGCAAAATTTG  
\*\*\*\*\*  
C18M AAGAAAAGAAATGAAC TAGAAATTATTGGAATTGGATAAATGGGCAAAATTTG  
\*\*\*\*\*  
C98H AAGAAAAGAAATGAACAAGACTTATTGGAATTAGATAAATGGGCAAGTTTG

NL43 TGGAAATTGGTTTAACATAACAAGAAATTGGCTGTGGTATATAAAATTATTCAT 8271  
\*\*\*\*\*  
D36P TGGAAATTGGTTTGACATAACAAGAAATGGCTGTGGTATATAAAATTATTCAT  
\*\*\*\*\*  
C18S TGGAAATTGGTTTAGTATATCAAACTGGCTATGGTATATAAAATTATTCAT  
\*\*\*\*\*  
C18M TGGAAATTGGTTTAGTATATCAAACTGGCTATGGTATATAAAATTATTCAT  
\*\*\*\*\*  
C98H TGGAAATTGGTTTGACATAACAAGTGGGCTGTGGTATATAAAATTATTCAT

4/47

## FIGURE 1

NL43 AATGATAGTAGGAGGCTTGGTAGGTTTAAGAAATAGTTTTTGCTGTACTTT 8321  
\*\*\* \*\*  
D36P AATGGTAGTAGGAGGCTTGTATAGGTTTAAGAAATAGTTTTTGCTGTACTTT  
\*\*\* \*\*  
C18S AATGGTAGTAGGAGGCTTGGTAGGTTTAAGAAATAGTTTTTACTGTACTTT  
\*\*\* \*\*  
C18M AATGGTAGTAGGAGGCTTGGTAGGTTTAAGAAATAGTTTTTACTGTACTTT  
\*\*\* \*\*  
C98H AATGATAGTAGGAGGCTTGGTAGGTTTAAGAAATAGTTTTTAGCTGTACTTT

SA8 SA9

SA10

NL43 CTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTAATCGTTTCAGACC 8371  
\*\*\* \*\*  
D36P CTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTAATCGTTTCAGACC  
\*\*\* \*\*  
C18S CTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTAATCGTTTCAGACC  
\*\*\* \*\*  
C18M CTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTAATCGTTTCAGACC  
\*\*\* \*\*  
C98H CTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTAATCGTTTCAGACC

5/47

FIGURE 1

Tat termination NL43

NL43 CACCTCCCAATCCGAGGGACCCGACAGGCCCGAAGGAATAGAAAGA 8421  
 \* \* \* \* \*  
 D36P CTCCTCCCAACCCGAGGGACCCGACAGGCCCGAAGGAATCGAAGAAGA  
 \* \* \* \* \*  
 C18S CACCTCCCAACCCGAGGGACCCGACAGGCCCGAAGGAATCGAAGAAGA  
 \* \* \* \* \*  
 C18M CACCTCCCAACCCGAGGGACCCGACAGGCCCGAAGGAATCGAAGAAGA  
 \* \* \* \* \*  
 C98H CACCTCCCAACCCGAGGGACCCGACAGGCCCGAAGGAATCGAAGAAGA

NL43 AGGTGGAGAGAGACAGAGACAGATCCATTGATTAGTGAACGGATCCT 8471  
 \* \* \* \* \*  
 D36P AGGTGGAGAGAGACAGAGACAGATCCACTCGATTAGTACACGGATTCT  
 \* \* \* \* \*  
 C18S AGGTGGAGAGAGACAGAGGAGCTCCACTCGATTAGTGCACGGATTCT  
 \* \* \* \* \*  
 C18M AGGTGGAGAGAGAGGAGGAGCTCCACTCGATTAGTGCACGGATTCT  
 \* \* \* \* \*  
 C98H AGGTGGAGAGAGACAGAGACAGATCCAGTCGATTAGTGCACGGATTCT  
 D36P, C18S, C18M & C98H Tat termination

6/47

## FIGURE 1

NL43	TAGCACTTATCTGGACGATCTGCGAGCCTGTGCCCTCTTCAGCTACCAC	8521
	*****	
D36P	TAGCACTTTTCTGGACGACCTGAGGAGCCTGTGCCCTCTTCCTCTACCAC	
	*****	
C18S	TAGCACTTTTCTGGACGACCTGAGGAGTCTGTGCCCTCTTCAGCTACCAC	
	*****	
C18M	TAGCACTTTTCTGGGTCGACCTGAGGAGTCTGTGCCCTCTTCAGCTACCAC	
	*****	
C98H	TAGCACTTTTCTGGGTCGACCTGAGGAGCCTGTGCCCTCTTCAGCTACCAC	
	*****	
NL43	CGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACCTTCTGGG	8571
	*****	
D36P	CACTTGAGAGACTTACTCTTGATTGTAACAGGATTGTGGAACCTTCTGGG	
	*****	
C18S	CACTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACCTTCTGGG	
	*****	
C18M	CGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACCTTCTGGG	
	*****	
C98H	CGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACCTTCTGGG	

7/47

## FIGURE 1

NL43 ACGCAGGGGTGGGAAGCCCTCAAATATTGGTGAATCTCCTACAGTATT 8621  
 \*\*\*\*\*  
 D36P ACGCAGGGGATGGGAAGCCCTCAAATATTGGTGAACCTCCTAAAGTATT  
 \*\*\*\*\*  
 C18S ACGCAGGGGATGGGAAGCCCTCAAATACTGGTGAATCTCCTGCAGTATT  
 \*\*\*\*\*  
 C18M ACGCGGGGATGGGAAGCCCTCAAATACTGGTGAATCTCCTGCAGTATT  
 \*\*\*\*\*  
 C98H ACGCAGGGGTGGGAAGCCCTCAAATATTGGTGAATCTCCTACAATATT

## NL43 Rev termination

NL43 GGAGTCAGGAACATAAGAAATAGTGCTGTTAAGTCTCAATGCCACAGCC 8671  
 \*\*\*\*\*  
 D36P GGAGCCAGGAACCTGCAGAAGAGTGCTGTATCTTGCTCAATGCCACCGCC  
 \*\*\*\*\*  
 C18S GGAGGCAGGAACCTACAGAAGAGTGCTGTAGCTTGTTCATGGCACGGCC  
 \*\*\*\*\*  
 C18M GGAGACAGGAACCTACAGAAGAGTGAGTTAGCTTGTTCATGCCATAGCC  
 \*\*\*\*\*  
 C98H GGAGTCAGGAACCTCAAGAAGAGTGCTATTAGCTTGTTCATGCCACCGCC  
 C18S, C18M & C98H Rev termination

8/47

## FIGURE 1

NL43 ATAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAGTATTACAAGCAGC 8721  
\*\*\*\*\*  
D36P ATAGCAGTAGCTGAGGGGACAGATAGAGTTTATAGAAGTATTACAAGAGC  
\*\*\*\*\*  
C18S ATAGCAGTAGCTGAGGGGACAGATAGAGTTATAGAAGCTTTACGAAGGCC  
\*\*\*\*\*  
C18M ATAGCAGTAGCTGAGGGGACAGATAGAGCTATAGAAGGATTACAAGAGC  
\*\*\*\*\*  
C98H ATAGCAGTAGCTGAGGGGACAGATAGAGTTATAGAAGTATTACAAGAGC  
\*\*\*\*\*  
D36P Rev termination

NL43 TTATAGAGCTATTCCGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAA 8771  
\*\*\*\*\*  
D36P TTATAGAGCTATCTCCACATACCTAGAAGAATAAGACAGGGCCTCGAAA  
\*\*\*\*\*  
C18S TTATAGAGCTATTCTCCACATACCTAGAAGAATAAGACAGGGCTTAGAAA  
\*\*\*\*\*  
C18M TTATAGAGCTATTCTCCACATACCTAGAAGAATAAGACAGGGCCTTAGAAA  
\*\*\*\*\*  
C98H TTGTAGAGCTGTTCTCCACATACCTAGAAGAATAAGACAGGGCTTCGAAA



9/47

FIGURE 1

	Env termination	Nef start	
NL43	GGATTTTGCTATAAAGATGGGTGGCAAGTGGTCAAAAGTAGTGATTGG	8821	
D36P	* ****	*****	*
	TGGCTTTGCTATAAAATGGGTGGCAAGTGAGCAAAAGTAGTGATCAG		
C18S	** ****	*****	** *
	GGGCTTTGCTATAAAATGGGTGGCAAGTGGTCAGAAAGTAGTGTTAG		
C18M	** ****	*****	
	GGGCTTTGCTGTAAATGGG-----		
C98H	** ****	***	
	GGGCTATGCTATAAAATGGGTGGCAAGTGGTTAAAGTAGTATGGTTAG		
	D36P Nef termination		
	nef duplication region		
NL43	ATGGCCCTGCTGTAAAGGGAAGAATGAGACGAGCTGAGCCAGCAGCATG	8871	
	** ** *	** ** *	*
D36P	ATAGCATG-----	-----CATCATAAG	
	* ** *		
C18S	AAGGCATG-----	-----	
C18M	-----	-----	
	*****	*****	*
C98H	ATGGCCCTGCTGTAAAGGGAATAATGAACAAGCTGAGCCAGCAGCAAG		

10/47

## FIGURE 1

NL43 GGGTGGGAGCAGTATCTCGAGACCCTAGAAAACATGGAGCAATCACAAGT 8921  
 \*\*\*\*\* \*\*  
 D36P GGGTGGGGGC-----  
 C18S -----  
 C18M -----  
 \*\*\*\*\*  
 C98H GGGTGGGAGCAATATCTCGAGACCCTAGGAAAACATGGAGCAATCCCAAGT

SIV<sub>mac</sub>239 IPTC

NL43 AGCAATACAGCAGCTAACAAATGCTGCTTGTGCTGGCTAGAGCACAAGA 8971  
 \*\* \*\* \*\*\*\*\* \* \*\*\*\*\*  
 D36P -----CAACAATAACAATGCTGATCGTGCTGGCTAGAGCACAAGA  
 C18S -----  
 C18M -----  
 \*\*\*\*\*  
 C98H AGCAATACAACAATAACAATGCTAATTGTGCTGGCTAGAGCACAAGA

11/47

FIGURE 1

NL43	GGAGGAAGAGGTGGGTTTCCAGTCACACCTCAGTACCTTTAAGACCAA	9021
	* * * * *	
D36P	GAAGGAAGAAGCGGGTTTCCAGTCAAACCTCAGTA-----	
	*****	
C18S	-----TACCTTTAAGAC-----	
C18M	-----	
	*****	
C98H	GGAGGAGGAAGTGGGTTTCCAGTCAAACCTCAGTACCTTTAAGACCAA	

**Poly purine tract**

NL43	TGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTAAAGAAAGGGG	9071
D36P	-----GCTGTAGATCTTAGCCACTTTTAAAGAAAGGGG	
C18S	-----AAGGCAGCTATAGATCTTAGCCGCTTTTAAAGAAAGGGG	
C18M	-----GATCTTAGCCACTTTTAAAGAAAGGGG	
C98H	TGACTTACAAG-----GCCACTTTTAAAGAAAGGGG	
C54P	AGCCACTTTTAAAGAAAGGGG	

C18S & C18F nef termination  
C18M and C98H nef Termination

12/47

FIGURE 1

	[ U3	
NL43	GGACTGGAAGGGCTAATTCACTCCCAAGAAGACAAGATATCCTTGATCT	9121
	*****	
D36P	GGACTGGAAGGGCTAATTCACTCCCAAGAAGACAAGATA-----	
	*****	
C18S	GGACTGGAAGGGCTAATTCACTCACAGAGAAGA-----	
	*****	
C18M	GGACTGGAAGGGCTAATTCACTCACAGAGAAGA-----	
	*****	
C98H	GGACTGGAAGGGCTAATTCACTCCCTAAAGAAGACAAGATATCCTTGATCT	
	*****	
C54P	GGACTGGAAGGGCTAATTCCGCTCCCAAGAAGACAAGATATCCTTGATCT	
	*****	
	SA12	
NL43	GTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTACACAC	9171
D36P	-----	
C18S	-----	
C18M	-----	
	*****	
C98H	TTGGATCTACCACACACAAGGCTACT-----	
	*****	
C54P	GTGGGTCTACCACACACAAGGCTACTTCCCTGAGTGGCAGAACTACACAC	

13/47

FIGURE 1

	[ NRE -->	
NL43	CAGGGCCAGGGTCCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTA	9221
D36P	-----*	*****
C18S	-----CACAGTGCTGCAAACTA	
C18M	-----	
C98H	-----ATCCACTGACTTTTGG, TGGTGCTTCAAAATTA	
C54P	CAGGGCCAGGACCAGATATCCACTGACCTTTGGATGGTGCTGCAAAACGA	
	myb	NF-AT
NL43	GTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCCAATAAAGGAGAGAACAC	9271
D36P	*****	*****
C18S	TTACCAGTGGAGTCAGCGAAGATAGAAGAGGCCCAATGGAGGAGAAACCA	
C18M	-----TCAGTTGAACCCAGAAAGATAGAAGAGGCCCATGAAGAAGAAACAA	
C98H	-----TCAGTTGAACCCAGAAAGATAGAAGAGGCCCATGAAGAAGAAACAA	
C54P	GTACCAGTGGANCCAGA--AGAGAGAAGAGACCCAATGGAGGAGAGAACA-	
	GTGCCAGTGGAAACAGAGAAGATAGAAGAGGCCCAATGGAGGAGAAACAA	
	(myb)	

14/47

FIGURE 1

	NF-AT	
NL43	<u>CAGCTTGTTACACCCGTGTGAGCCTGCATGGAATGGATGACCCCTGAGAGAG</u>	9321
	*** ****	
D36P	CAGATTGTT-----	
	*** ****	
C18S	CAGATTGTT-----	
	*** ****	
C18M	CAGATTGTT-----	
	***	
C98H	-----	
	*** ****	
C54P	CAGACTGTT-----	
	-----	
	USF	
	<-- NRE ]	
NL43	<u>AAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCAGGTGGCC</u>	9371
	-----	
D36P	-----	
	* * *	
C18S	-----CCGTTTGTT	
	***	
C18M	-----CTGCT	
	-----	
C98H	-----A	
	-----	
C54P	-----	

15/47

FIGURE 1

TCF-1 $\alpha$ 

NL43	CGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCTA	Nef termination	9421
	* * * * *		
D36P	CTGTTGGGGACTTTCCATCCGTTGGGGACTTTCCAAAGCGCGGTGGCCTG		
	* * * * *		
C18S	CCGTTGGGGACTTTCCA,,,,,GGAGACGTGGCCTGAGTGATAAGCCG		
	* * * * *		
C18M	TGCTCAGCTGGGGACTTTCCAGAAAGCGCGCCTGAGTGACTAAGCCCCG		
	* * * * *		
C98H	CAGAGTGTGGGGACTCTCCACAACAGAGTGTGGGGACTTTCCAAAGGAGGC		
	* * * * *		
C54P	-----CCGTTGGGGACTTTCCAAGGAGCGCGTGGCCTGAGTGACTAAGTTCC		

D36P, C18S, C18M, &amp; C98H extra NFKB

D36P &amp; C98H extra NFKB





17/47

FIGURE 1

	Sp1	Sp1	TATA box	
NL43	CTGGCGGGGACTGGGAGTGGCGAGCCC, TCAGATGCTGCATATAAGCAG			9510
D36P	CTGGCGGGGACTGGGAGTGGCGAGCCC, TCAGATGCTGCATATAAGCAG			
C18S	CTGGCGGGGACTGGGAGTGGCGAGCCC, TCAGATGCTGCATATAAGCAG			
C18M	CTGGCGGGGACTGGGAGTGGCGAGCCC, TCAGATGCTGCATATAAGCAG			
C98H	CTGGCGGGA-CTGGGAGTGC-GAGCC-, TCAGATGCTGCATATAAGCAG			
C54P	CTGGCGGGGACTGGGAGGGGCGAGCCC, TCAGATGCTGCATATAAGCAG	Sp1		
	U3	R	TAR	
NL43	CTGCTTTTTCCTGTACTGGGTCTCTCTGGTTAGACCCAGATCTGAGCCTG			9560
D36P	CTGCTTTTTCCTGTACTGGGTCTCTCGGGTTAGACCCAGATCTGAGCCTG			
C18S	CTGCTTTTTCCTGTACTGGGTCTCTCTGGTTAGACCCAGATCTGAGCCTG			
C18M	CTGCTTTTTCCTGTACTGGGTCTCTCTGGTTAGACCCAGATCTGAGCCTG			
C98H	CTGCTTTTTCCTGTACTGGGTCTCTCTGGTTAGACCCAGATCTGAGCCTG			
C54P	CTGCTTTTTCCTGTACTGGGTCTCTCTGGTTAGACCCAGATCTGAGCCTG			

18/47

FIGURE 1

	Polyadenylation	
NL43	<u>GGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCCTCAATAAAGCT</u> *****	9610
D36P	<u>GGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCCTCAATAAAGCT</u> *****	
C18S	<u>GGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCCTCAATAAAGCT</u> **	
C18M	<u>GGAGCTCTCTGGCTAGTAGGGACCCACTCCTTAAGCCCTCAATAAAGCT</u> *****	
C98H	<u>GGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCCTCAATAAAGCT</u> *	
C54P	G incomplete	
R ] [ U5		
NL43	<u>TGCCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT</u> *****	9660
D36P	<u>TGCCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT</u> *****	
C18S	<u>TGCCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT</u> *****	
C18M	<u>TGCCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT</u> *****	
C98H	<u>TGCCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT</u> *****	

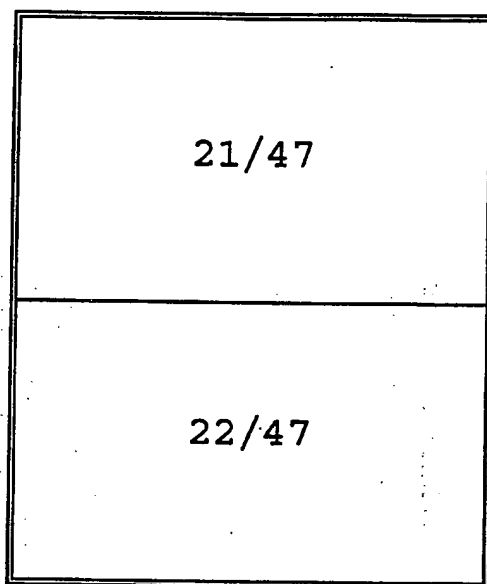
19/47

## FIGURE 1

NL43	AACTAGAGATCCCCTCAGACCCCTTTTAGTCAGTGTGGAAAATCTCTAGCA	9709
	* * * * *	
D36P	ATCTAGA	1305
	* * * * *	
C18S	ATCTAGA	1209
	* * * * *	
C18M	ATCTAGAGATCCCCTCAGACCCATTTTAGTCCGTGTGGAAAATCTCTAGCA	END
	* * * * *	
C98H	ATCTAGA	1399

20/47

FIGURE 2



21/47

## FIGURE 2A

NL43	73	PTSQSRGDPTGPKE#	86
D36P BMC		PSSQPRGDPTGPKEKKKVERETETDPLD#	
C18 HIV <sub>Stv</sub>		PTSQPRRDPTGQESKKKVERETEAAPLD#	
C18 HIV <sub>MBC</sub>		PTSQPRRDPTGQESKKKVERETEAAPLD#	
C98 HIV		PTSQPRRDPTGQESKKKVERETETDPVD#	

## FIGURE 2B

NL43	26	DPPNPEGTRQARRNRRRRWRERQRQIHSISERILSTYLG	65
D36P BMC		DPPNPEGTRQARRNRRRRWRERQRQIHSISTRILSTFLG.	
C18 HIV <sub>Stv</sub>		DPPNPEGTRQARRNRRRRWRERQRQLHSISARILSTFLG	
C18 HIV <sub>MBC</sub>		DPPNPEGTRQARRNRRRRWRERQRQLHSISARILSTFLG	
C98 HIV		DPPNPEGTRQARRNRRRRWRERQRQIQSISARILSTFLG	

22/47

## FIGURE 2B

NL43	RS AEPVPLQLPPLERLTLD CNEDCGTSGTQGVGSPQILVE	105
D36 PBMC	RPEEPVPLPLPPLERLTLD CNKDCGTSGTQGMGSPQILVE	
C18 HIV <sub>Stv</sub>	RPEESVPLQLPPLERLTLD CNEDCGTSGTQGMGSPQILVE	
C18 HIV <sub>MBC</sub>	RPEESVPLQLPPLERLTLD CNEDCGTSGTQGMGSPQILVE	
C98 HIV	RPEEPVPLQLPPLERLTLD CNEDCGTSGTQGVGSPQILVE	
NL43	SPTVLESGTKE#	116
D36PBMC	PPKVLEPGTAEECCYLAQCHRHSSS#	
C18 HIV <sub>Stv</sub>	SPAVLEAGTTEEC#	
C18 HIV <sub>MBC</sub>	SPAVLEAGTTEEC#	
C98 HIV	SPTILESGTQEECY#	

23/47

FIGURE 3

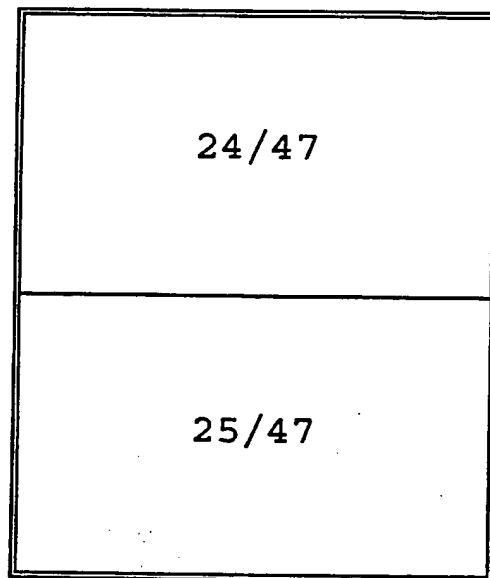


FIGURE 3

639

NL43  
D36P BMC  
C18 HIV<sub>Stv</sub>  
C18 HIV<sub>MBC</sub>  
C98 HIV

EQIWNMTWMEWDREINNNTSLIHSLIEESQSQEKNEQELLELDKWASL  
EEIWENMTWMQWEKEIHNHTKYIYSLLEKSNQSQEKNEQELLELDQWASL  
ETIWDNMTWMQWEREIDNYTNIYTLIEESQSQEKNELELELDKWANL  
EINNYTRTIYNLIEESQSQEKNEQDLELDKWASL

689

NL43  
D36 P BMC  
C18 HIV<sub>Stv</sub>  
C18 HIV<sub>MBC</sub>  
C98 HIV

WNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQT  
WNWFEDITKWLWYIKLFIMVVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQT  
WNWFSISNWLWYIKLFIMVVGGLVGLRIVFTVLSIVNRVRQGYSPLSFQT  
WNWFEDITSGWLWYIKLFIMIVGGLVGLRIVLAVLSIVNRVRQGYSPLSFQT



25/47

## FIGURE 3

NL43	HLPIDRGPDPEGIEEGGERDRDRSIRLVNGSLALIWDDLRLSLCLFSYH	739
D36 PBMC	LLPTPRGPDPEGIEEGGERDRDRSTRLVHGFLALFWDDLRLSLCLFLYH	
C18 HIV <sub>StV</sub>	HLPTPKGPDPEGIEEGGERDRGSSTRLVHGFLALFWDDLRLSLCLFSYH	
C18 HIV <sub>MB</sub>		
C98 HIV	HLPTPRGPDPEGIEEGGERDRDRSSRLVHGFLALFWVDLRLSLCLFSYH	
NL43	RLRDLLLIVTRIVELLGRRGWEALKYWNLLQYWSQELKNSAVNLLNATA	789
D36 PBMC	HLRDLLLIVTRIVELLGRRGWEALKYWNLLKYWSQELQKSAVILLNATA	
C18 HIV <sub>StV</sub>	HLRDLLLIVTRIVELLGRRGWEALKYWNLLQYWRQELQKSAVSLFNGTA	
C18 HIV <sub>MB</sub>		
C98 HIV	RLRDLLLIVTRIVELLGRRGWEALKYWNLLQYWSQELKKS AISLFNATA	
NL43	IAVAEGTDRVIEVLQAAVRAIRHIPRRIRQGLERILL#	839
D36 PBMC	IAVAEGTDRVLEVLQRAYRAILHIPRRIRQGLEMALL#	
C18 HIV <sub>StV</sub>	IAVAEGTDRVIEALRRAYRAILHIPRRIRQGLERALL#	
C18 HIV <sub>MB</sub>		
C98 HIV	IAVAEGTDRVIEVLQACRAVLHIPRRIRQGFERAML#	

26/47

FIGURE 4

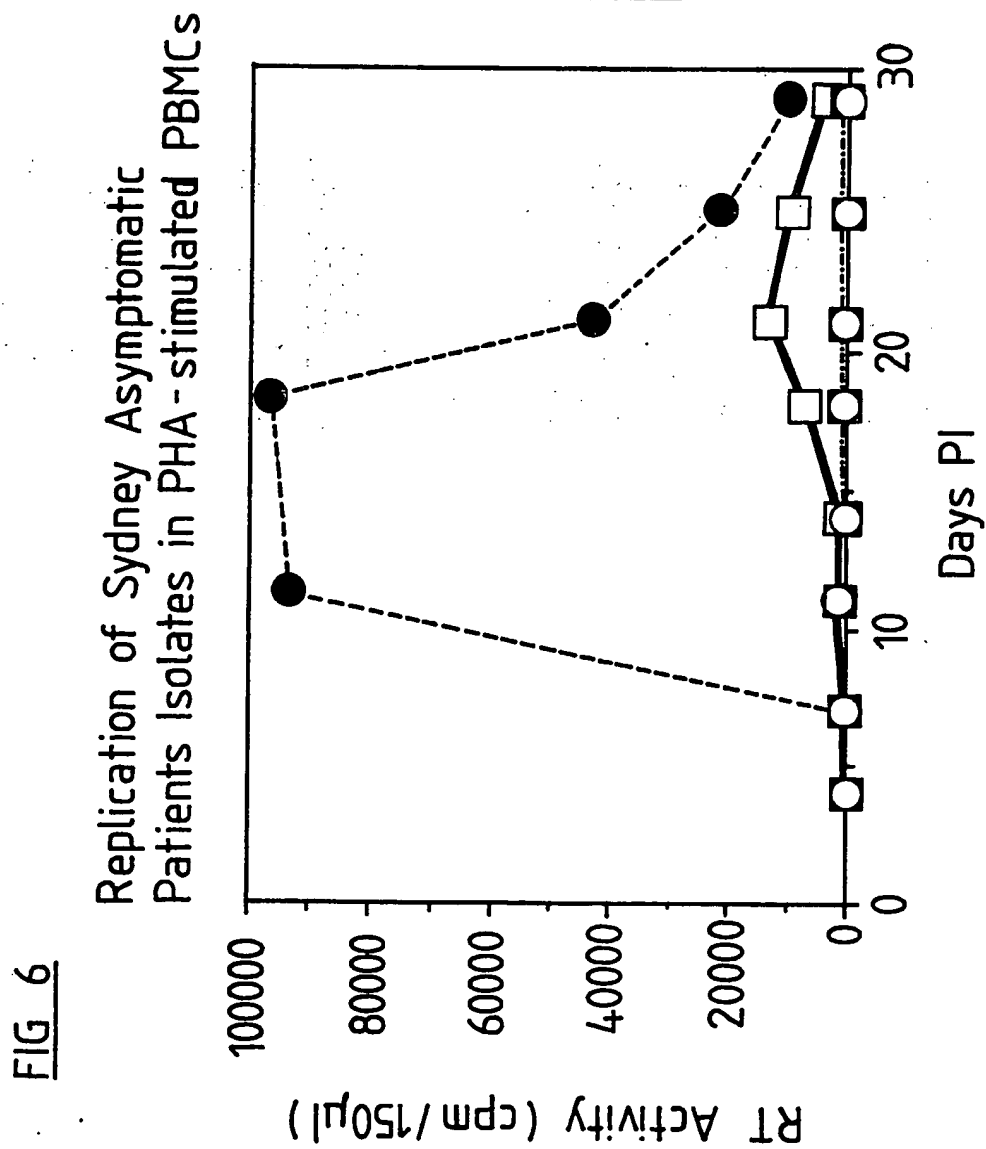
NL43	MGGKWSKSSVIGWPAVRERMRRRAEPAADGVGAVSRDLEKHGAITSSNTAA ***	50
D36 PBMC	MGGK# *****	4
C18.HIV <sub>stV</sub>	MGGKWSSESSVRRHVPLRQGSYRS# *	24
C18 HIV <sub>MC</sub>	MRILATF# *****	7
C98 HIV	MGGKWLKSSMVRWPAVREKMKQAEPAEGVGAISRDLGKHGAIPSSNTTT *****	50
NL43	NNAACAWLEAQEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGL ***	100
C98 HIV	NNANCAWLEAQEEEEVGFPVKPQVPLRPMTYKATF# *****	85
NL43	IHSQRRQDILDLWIYHTQGYFPDWNQNYTPGPGVRYPLTFGWICYKLVPVEP	150
NL43	DKVEEANKGENTSLLHPVSLHGMDDPEREVLEWRFFDSRLAFHHVARELHP	200
NL43	EYFKNC*	206

27/47

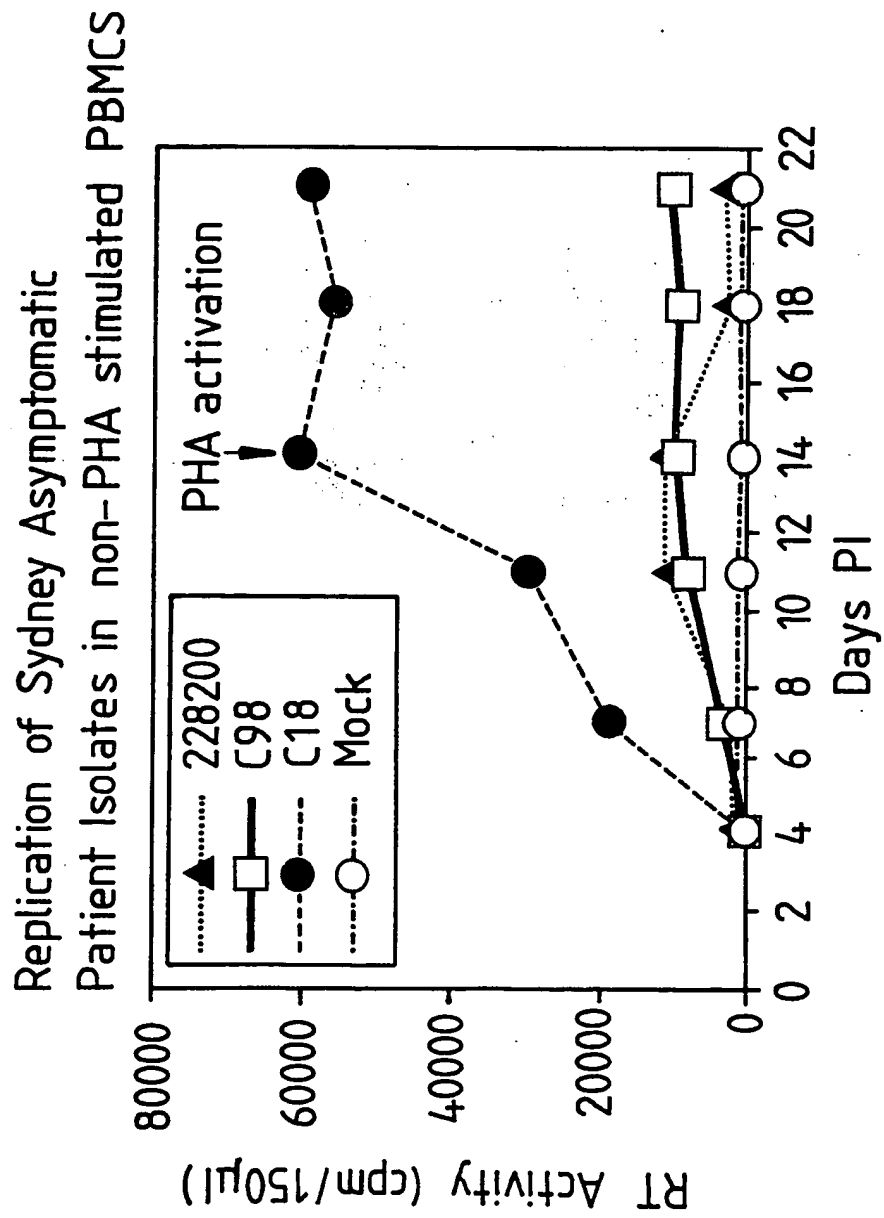
FIGURE 5

	NFkB	NFkB
9419 NL43	CGAGCTTGCTACAAGGGACTTTCC,,GCTGGGACTTTCCAGGGA	
D36 PBMC	ACTGTTGGGACTTTCCATCCGTGGGACTTTCCAAAGGC	
C18 HIV <sub>Stv</sub>	ΔCCGTTTGTTCCGTTGGGACTTTCCA-GGA	
C18 HIV <sub>MBC</sub>	ΔCTGCTTGTCTCAGCTGGGACTTTCCA-GAA	
C98 HIV	ΔACAGAGTGTGGGACTTCTCCACAACAGAGTGTGGGACTTTCCAAAGGA	
C54 PBMC	ΔCCGTTGGGACTTTCCAAAGGA	
	NFkB	NFkB
	Sp1	Sp1
NL43	GGCGTGGCCCTGGCGGACTGGGGAGTGGCG-AGCCCTCA	9492
DC36 PBMC	GGCGTGGCCCTGGGTGACTAGTTCGGTGGGG-ACTTTCCA	
C18 HIV <sub>Stv</sub>	GACGTGGCCCTGAGTGACTAAG-CCGCTGGGG-ACTTTCCG	
C18 HIV <sub>MBC</sub>	GGCGGGCCCTGAGTGACTAAGCCCCCGTTGGG-ACTTTCCG	
C98 HIV	GGCGTGGCCCTGAGTGACTAAGTTCCGTTGGGACTTTCCA	
C54 PBMC	GGCGTGGCCCTGAGTGACTAAGTTCCGTTGGGACTTTCCA	
	Sp1	3' half NFkB
		NFkB

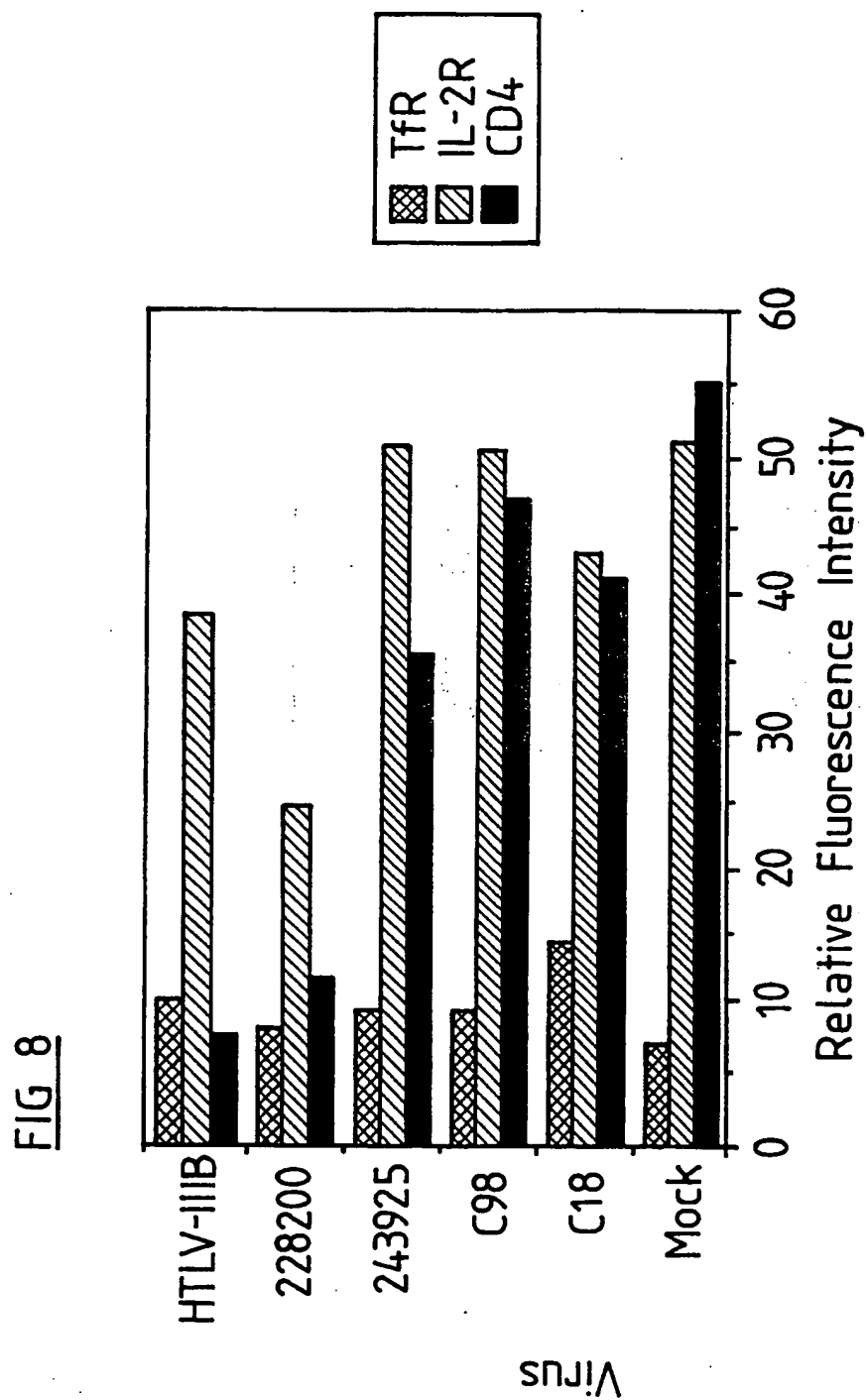
28/47



29/47

**FIG 7**

30/47



31/47

FIGURE 9

32/47
33/47
34/47
35/47
36/47
37/47

FIGURE 9

32/47

TGGAAGGGCTAATTCACCTCACGGAAAAGACCAGTTGAACCAG  
AAGAAGATAGAAGAGGGCCATGAAGAAGAAAACAACAGATTGT  
TCTGCTTGCTCAGCTGGGGACTTTCCAGAAGGCGCGGCCTGA  
GTGACTAAGCCCCGTTGGGGACTTTCCGAAGAGGCATGAAGG  
GACTTTCCAAGGCAGGCGTGGCCTGGGCGGGACTGGGGAGTG  
GCGAGCCCTCAGATGCTGCATATAAGCAGCTGCTTTCTGCCT  
GTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGC  
TCTCTGGCTAGCTAGGGAACCCACTGCTTAAGCCTCAATAAA  
GCTTGCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGT  
GTGACTCTGGTATCTAGAGATCCCTCAGACCATTTTAGTCCG  
TGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAA  
AGCGAAAGGAAAACCAGAGGAGCTCTCTCGACGCAGGACTCG  
GCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACT  
GGTGAGTACGCCGAAAATTTTGACTAGCGGAGGCTAGAAGGA  
GAGAGATGGGTGCGAGAGCGTCAATATTAAGCGGGGGAAAAT  
TAGATAGATGGGAGAAAATTCGGTTAAGGCCAGGAGGAAAGA  
AAAAGTATAAATTAACACATATAGTATGGGCAAGCAGGGAGC  
TAGAACGATTCGCAGTCAATCCTGGCCTGTTGGAAACATCAG  
AAGGCTGTAGACAAATACTGGGACAGTTACACCCGTCCCTTC  
AGACAGGATCAGAAGAACTTAAATCAGTATATAATGCAGTAG  
CAGTCCTCTATTGTGTGCATCAAAACATAGACATAAAGGACA  
CCAAGGAAGCTTTAGAAAAGATAGAGGAAGAGCAAAACAAAT  
GTAAGAAAAAAGCACAGCAAGCAGCAGCACAGCAAGCAGCAG  
CTGGCACAGGAAACAGCAACCCGGTCAGCCAAAATTACCCTA  
TAGTACAGAACATGCAGGGGCAAATGGTACATCAGGCCATAT  
CACCTAGAACTTTAAATGCATGGGTAAAAGTAATAGAAGAGA  
AGGCTTTCAGCCCAGAGGTAATACCCATGTTTTTCAGCATTAT  
CAGAAGGAGCCACCCACAAGATTTAAACACCATGCTAAACA  
CAGTGGGGGGACATCAAGCAGCTATGCAAATGTTAAAGAGA  
CCATCAATGAGGAAGCTGCAGAATGGGATAGATTACATCCAG  
CGCAGGCAGGGCCTGTTGCACCAGGCCAGATGAGAGACCCAA  
GGGAAGTGACATAGCAGGAAC TACTAGTACCCTTCAGGAAC  
AAATAGGATGGATGACAGGTAATCCAGCTATCCCAGTAGGAG  
AAATCTATAAAAGATGGATAATCCTGGGATTAAATAAAATAG  
TAAGGATGTATAGCCCTATCAGCATTCTGGACATAAAACAAG  
GACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATA  
AACTCTAAGAGCCGAGCAAGCTACACAGGAGGTAAAAAATT  
GGATGACAGAAACCTTGTTGGTCCAAAATGCAAACCCAGATT  
GTAAGACTATTTTAAAGCATTTGGGACCAGCAGCTACACTAG



FIGURE 9

33/47

AAGAAATGATGACAGCATGTCAGGGAGTGGGAGGACCCAGCC  
ATAAAGCAAGAGTTTTTGGCAGAAGCAATGAGCCAAGCAACAA  
ATGCAGCTACTGTAATGATGCAGAGAAGCAATTTTAGAAACC  
AAAGAAAGAATGTTAAGTGTTTCAATTGTGGCAAAGAAGGGC  
ACATAGCCAGAAATTGCAGGGCTCCTAGGAAAAGGGGCTGTT  
GGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTACTG  
AGAGACAGGCTAATTTTTTTAGGGAAAATCTGGCCTTCCCACA  
AGGGGAGGCCAGGGAACTTTCTTCAGAGCAGGCCAGAACCAA  
CAGCCCCCTCTCCAGGGCAGGCCGGAGCCATCAGCCCCGCCAG  
AAGAGAGCTTCAGGTTTGGGGAGGAGACAACAACCTCCCTCTC  
AGAAGCAGGAGCCGATAGACAGGGACAGGGATCTGTATCCTT  
TAGCTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCAC  
AATAAAGATAGGGGGGCAGCTGAAGGAAGCTCTATTAGATAC  
AGGAGCAGATGATACAGTATTAGAAGACATGCATTTGCCAGG  
AAAATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTAT  
CAAAGTAAACAATATGATGAAATTCTTGTAAGAAATCTGTGG  
ACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGT  
CAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCAC  
TTTAAATTTTCCCATTAGTCTTATTGAACTGTACCAGTACA  
ATTAAAGCCAGGAATGGATGGCCCAAAGGTAAACAATGGCC  
ATTGACAGAAGAGAAAATAAAAGCATTAGTAGAAATTTGTAC  
AGAAATGGAAAAGGAAGGAAAGATTTCAAAAATTGGGCCTGA  
AAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGA  
TGGTACTAAATGGAGAAAATTAGTAGATTTTCAGAGACCTTAA  
TAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACC  
ACATCCCTCAGGATTAAAAAAGAAAAAATCAGTAACAGTACT  
GGATGTGGGTGATGCATACTTTTCAGTTCCCTTAGATGAAAA  
CTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAATAA  
TGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACA  
GGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGAC  
AGAATCTTAGAGCCTTTTAGAAGACAAAATCCAGACATAGT  
TATCTATCAATACATGGATGACTTGTATGTAGGATCTGATTT  
AGAAATAGGACAGCATAGAATAAAAATAGAGGAACTGAGACA  
ACATCTGTTGAAGTGGGGATTTACCACACCAGACAAAAGCA  
TCAGAAAGAACCCCATTCCTTTGGATGGGTATGAACTCCA  
TCCTGATAAATGGACAGTGCAACCTATAGTACTGCCAGAAAA  
AGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGTAA  
ATTAAATTGGGCAAGTCAGATTTACCCAGGAATTAAAGTAAG  
GCAATTATGTAAACTCCTTAGGGGAACCAAAGCACTAACAGA

FIGURE 9

34/47

AGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGA  
AAACAGGGAAATTCTAAGAGAACCAGTACATGGAGTGTATTA  
TGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGA  
GCAAGGCCAATGGACATATCAAATTTATCAAGATCAATTTAA  
AAATCTAAAAACAGGAAAGTATGCAAGATTGAGGGGTGCCCA  
CACTAATGATGTAAAACAATTTCCAGAGGCAGTGCAAAAAAT  
AGCCACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATT  
TAGACTACCCATACAAAAAGAAACATGGGACGCATGGTGGAC  
AGAGTATTGGCAAGCCACCTGGATTCTCTGAGTGGGAGTTTGT  
CAATACCCCTCCCCTAGTAAAATTATGGTACCAGTTAGAAAA  
AGAACCCATAATAGGAGCAGAACTTTCTATGTAGATGGGGC  
AGCTAACAGAGAGACTAAATTAGGAAAAGCAGGATATGTTAC  
TGACAGAGGAAGACAAAAAGTTGTCTCCCTAACTGACACAAC  
AAATCAGAAGACTGAGTTACAAGCAATTCATCTAGCTTTGCA  
GGATTCAGGATTAGAAGTAAACATAGTAACAGACTCACAGTA  
TGCATTAGGAATCATTCAAGCACAACCAGATAAAAGTGAATC  
AGAAATAGTCAATCAAATAATAGAGCAATTAATAAAAAAGGA  
AAAGGTCTACCTGGCATGGGTACCAGCACACAAAGGAATTGG  
AGGGAATGAACAAGTAGATAAATTAGTCAGTGCTGGAATCAG  
GAAAATACTATTTTTTAGATGGAATAGATAAGGCACAAGAAGG  
CCATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGG  
TTTTAACCTGCCACCTATAGTAGCAAAAGAAATAGTAGCCAG  
CTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACA  
AGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACA  
TCTAGAAGGAAAAATTATCCTGGTAGCAGTTCATGTAGCCAG  
TGGATATATAGAAGCAGAAGTTATTCCAGCAGAGACAGGGCA  
GGAAACAGCATACTTTATCTTAAATTAGCAGGAAGGTGGCC  
AGTAAACACAATACATACAGACAATGGCGGCAATTTTCATCAG  
TACCACGGTTAAGGCCGCCTGTTGGTGGGCAGGGATCAAGCA  
GGAATTTGGCATTCCCTACAATCCCCAAAGCCAAGGAGTAGT  
GGAATCTATGAATAGAGAATTAAAGAAAATTATAGGACAGGT  
AAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAATGGC  
AGTATTCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGG  
ATACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGA  
CATACAACTAAAGAATTACAAAAGCAAATTACAAAAATTCA  
AAATTTTCGGGTTTATTACAGGGACAGCAGAGATCCACTTTG  
GAAAGGACCAGCAAACTTCTCTGGAAAGGCGAAGGGGCAGT  
AGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGAAG  
AAAAGTAAAGATCATTAGGGATTATGGAAAACAGATGGCAGG

FIGURE 9

35/47

TGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGAACAT  
GGAACAGTTTGTAGTGAAACACCATATGTATGTTTCAAAGAAAG  
CTAAGGGATGGATTTATAGACATCACTATGAAAACACTCATC  
CAAAAATAAGCTCAGAAGTACACATCCCACTAGGGGAAGCTA  
GATTGGTAATAACAACATATTGGGGTCTACATACAGGAGAAA  
GAGACTGGCATTGTTGGGTGAGGGAGTCTCCATAGAATGGAGGG  
AAAGGACATATAGAACACAAGTAGACCCCGAACTAGCAGACC  
AACTAATTCATATACATTACTTTGATTGTTTTTTCAGAATCTG  
CCATAAGAAGTGCCATATTAGGATATAGAGTTAGGCATAGGT  
GTGAATATCAAGCAGGACATAACAAGGTAGGATCTCTACAGT  
ACTTGGCACTAACAGCATTAATAACACCAAAGAAGATAAAGC  
CACCTTTGCCTAGTGTGCGAAACTGACAGAGGATAGATGGA  
ACAAGCCCCAGAAGACCAAGGGCCACAGAGGCAGCCATACAA  
TGAATGGACACTAGAACTTTTAGAGGAGCTTAAGAATGAAGC  
TGTTAGGCATTTTCCTAGGGTATGGCTCCATGGCTTAGGGCA  
ACATATCTATGAACTTATGGGGTACTTGGGAAGGAGTGGA  
GGCCATAACAAGAACTCTGCAACAACCTGCTGTTTATTCAATT  
CAGAATTGGGTGTCAACATAGCAGAATAGGCATTATTTCGACA  
GAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAGACTAGAGC  
CCTGGAAGCATCCAGGAAGTCAGCCTAAGACTGCGTGTACCA  
CTTGCTATTGTAAAAAGTGCTGCTTTTCATTGCCAAGTTTGTT  
TTATGACAAAAGGCTTAGGCATCTCCTATGGCAGGAAGAAGC  
GGAGACAGCGACGAAGAGCTCCTCAAGACAGTCAGACTCATC  
AAGCTTATCTATCAAAGCAGTAAGTAATATATGTAATGCAAC  
CTTTACAAATAGTAGCAATAGTAGCATTAGTAGTAGCAGGAA  
TAATAGCAATAGTTGTGTGGACCATAGTATTCATAGAATATA  
AGAAAATATTAAGACAAAGAAAAATAGACAGGTTGATTGATA  
GAATAAGAGAAAGAGCAGAAGACAGTGGCAATGACAGTGAAG  
GGGATCAGGAAGAATTATCGGCACTTGTGGACATGGGGCACC  
ATGATCCTTGGGATATTAATGATCTGTAGAGCTGCAAACAAT  
TTGTGGGTACAGTCTATTATGGGGTACCTGTGTGGAGAGAA  
GCAACCACCACTCTATTTTGTGCATCAGATGCCAAGGCATAT  
GATGCAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTA  
CCCACAGACCCTAACCACACAAGAAGTAGAATTGAAAAATGTG  
ACAGAAAATTTTAACATGTGGAAAAATAACATGGTAGAACAG  
ATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTGAAG  
CCATGTGTAAAATTAACCCCACTCTGTGTTTCTTTAAATTGC  
ACTGATGCTACTAATAACCACTAATAGTAATAACCACTAGCAGC  
AGCGAGAAACCGAAGGGGACAGGGGAAATAAAAACTGCTCT

FIGURE 9

36/47

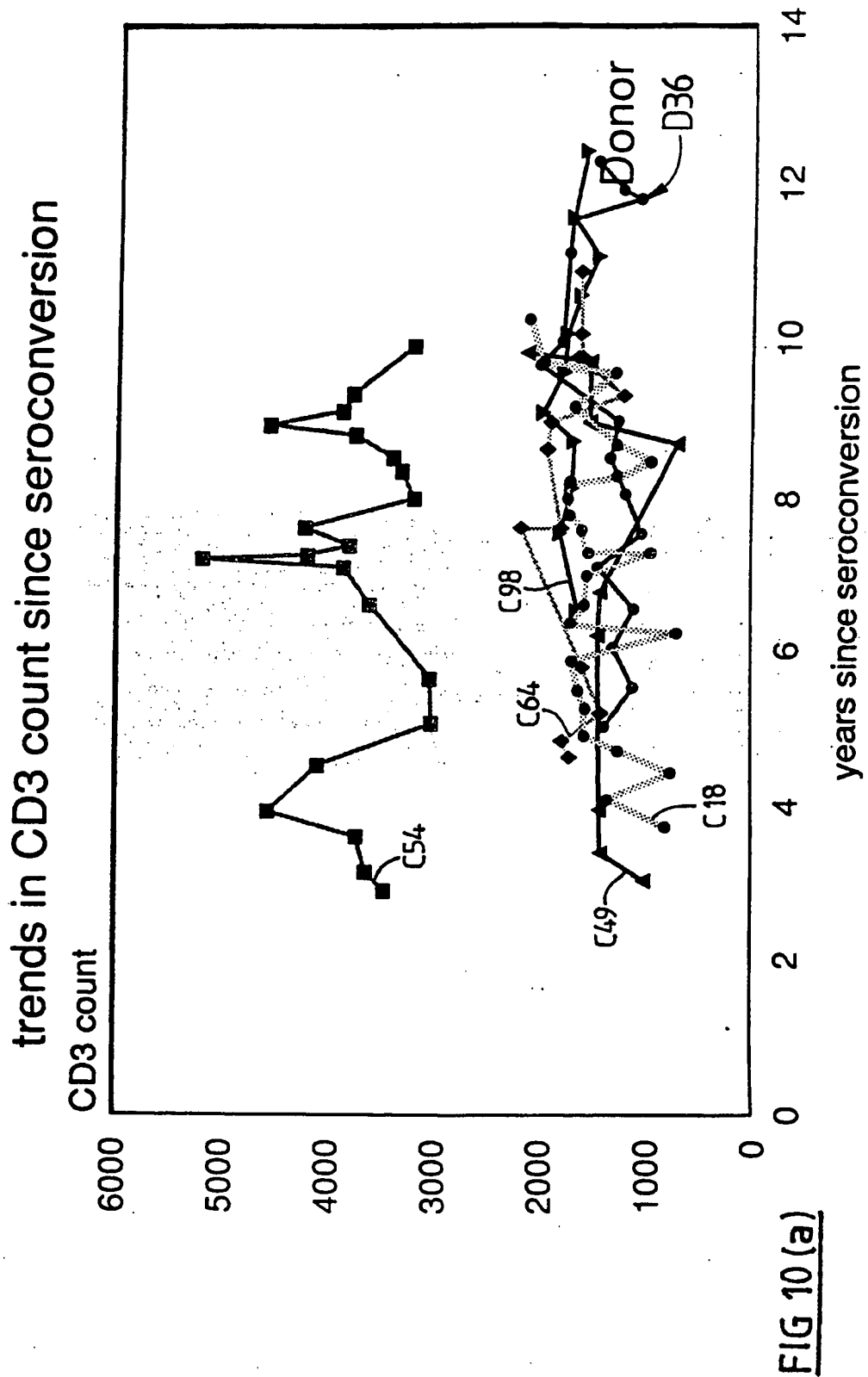
TTCAATATCACCACAAGCATAAGAGATAAGGTGCAGAAACAA  
TATGCACTTTTTTATAGCCTTGATGTAGTACCAATGGATGAT  
AATGATAATAGTACAAGCTATAGGTTAATAAGTTGTAACACC  
TCAATCATTACACAGGCCTGTCCAAAGATATCCTTTGAGCCA  
ATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTA  
AAGTGTAAGATAAAAGGTTCAATGGAAAAGGACCATGTACA  
AGTGTCAGCACAGTACAGTGTACACATGGAATTAGGCCAGTA  
GTATCAACTCAACTGTTGTTAAATGGCAGTCTAGCAGAAGAA  
GAGGTAGTAATTAGATCTGACAATTTTACGAACAATGCTAAA  
ACCATAATAGTACAGCTGAGCAAATCTGTAGAAATTACTTGT  
GTAAGACCCAACAACAATAACAAGAAAAAGTATAAGTATGGGA  
CCAGGGAGAGCATTTTTATACAACAGAAATAATAGGAGATATA  
AGACAAGCATATTGTAACATTAGTAAAGCAAACCTGGACTGAC  
ACTTTAGAACAGATAGCTAGAAAATTAAGAGAACAATTTGAG  
AATAAAACAATAGTCTTTAAGCCATCCTCAGGAGGGGACCCA  
GAAATTGTAACACAGTTTTTACAGTTTTTAATTGTGGAGGGGAA  
TTTTTCTACTGTAATTCAACACAACCTGTTTAATGGTACTTGG  
AATGGTACTTGGGTTAATGGTACTTGGAGTAGTAATAATACG  
ACTGATACTGCAAATATCACACTCCCATGCAGAATAAAACAA  
TTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCC  
CCTCCCATCAAAGGACAAATTAATGTACATCAAATATTACA  
GGGCTGATATTAACAAGAGATGGTGGTAACAATAACACCACG  
AACGACAACGAGACCGAGACCTTCAGACCTGGAGGAGGAGAT  
ATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTA  
GTACAAGTTGAACCATTAGGAGTAGCACCCACCAAGGCAAAG  
AGAAGAGTGGTGCAAAGAGAGAAAAAGAGCAGTGGGAATAGGA  
GCTATGTTCTTGGGTTCTTAGGAGCAGCAGGAAGCACTATG  
GGCGCAGCGTCAGTGACGCTGACGGTACAAGCCAGACAATTA  
TTGTCTGGTATAGTGCAGCAGCAGAACAATCTGCTGAGGGCT  
ATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGC  
ATCAAACAGCTCCAGGCAAGAGTCCTGGCTGTGGAAAGATAC  
CTAAGGGATCAACAGCTCCTGGGACTTTGGGGTTGCTCTGGA  
AACTCATTGTGACCACTACTGTGCCTTGGAACAATAGCTGG  
AGTAATAAATCTCTGGAAACAATTTGGGATAACATGACCTGG  
ATGCAGTGGGAAAGAGAAATTGACAATTACACAAACATAATA  
TACACCTTAATTGAAGAATCGCAGAACCAACAAGAAAAAAT  
GAACTAGAATTATTGGAATTGGATAAATGGGCAAATTTGTGG  
AATTGGTTTTAGTATATCAAACCTGGCTATGGTATATAAAATTA  
TTCATAATGGTAGTAGGAGGCTTGGTAGGTTTAAGAATAGTT

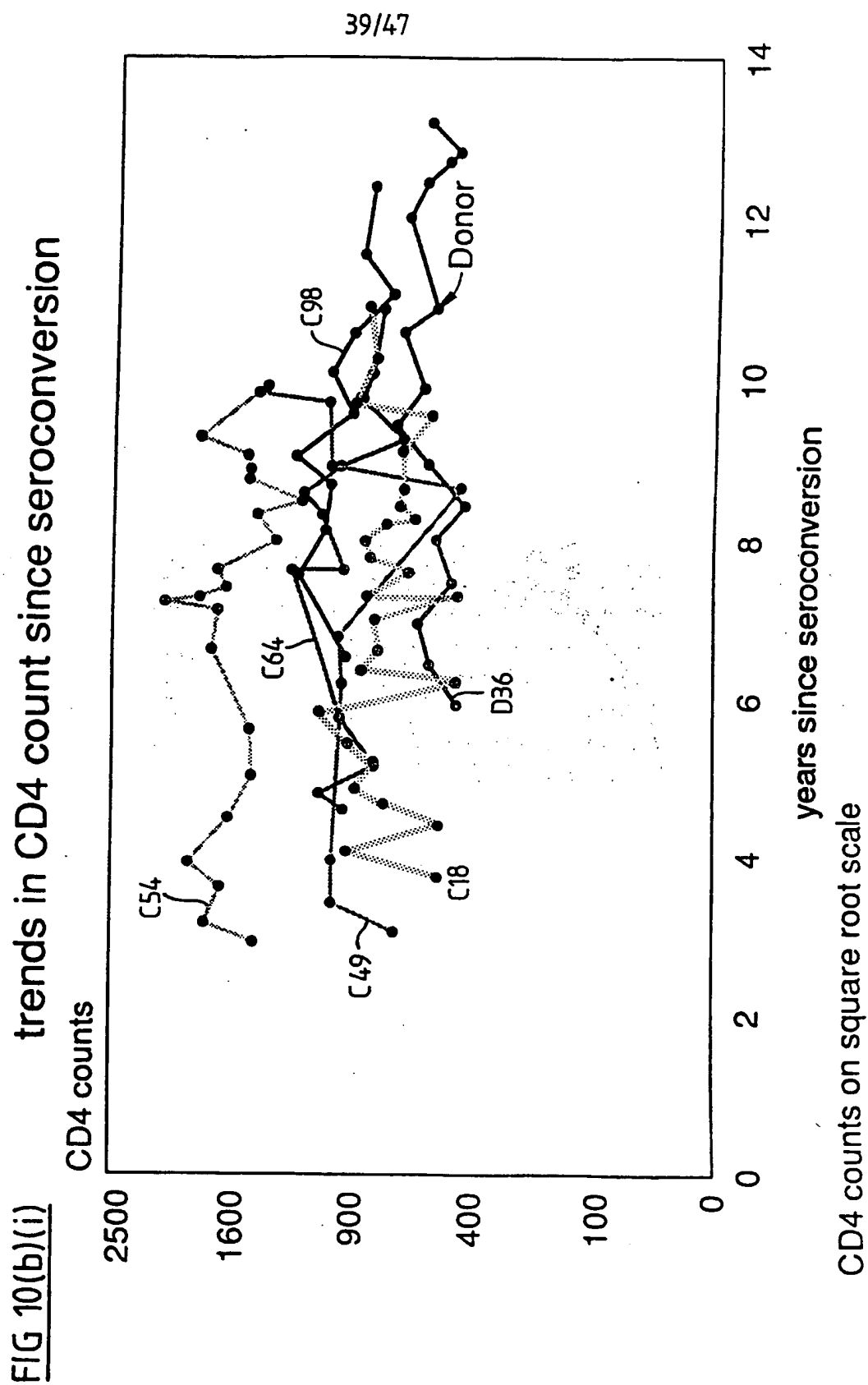
FIGURE 9

37/47

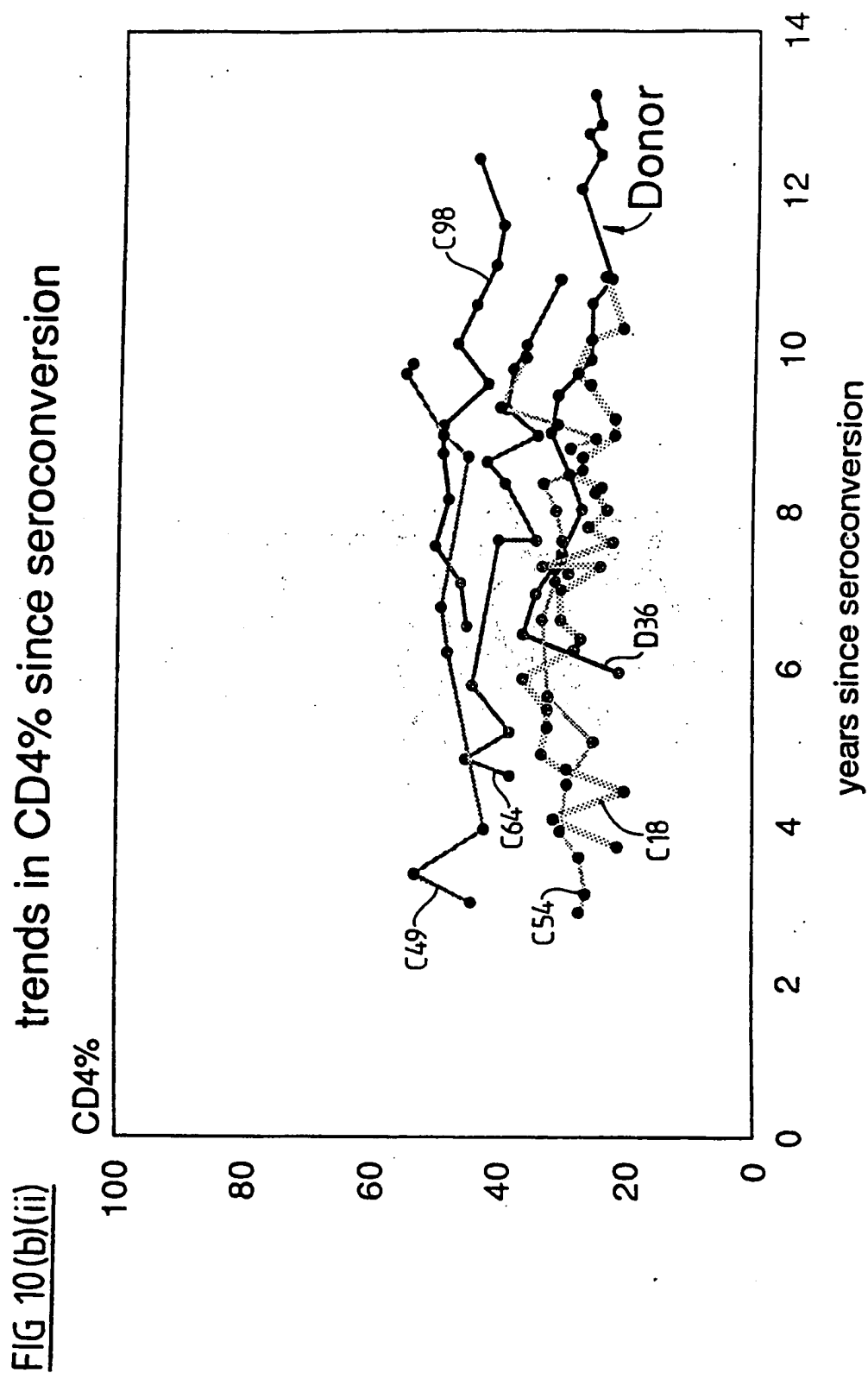
TTTACTGTACTTTCTATAGTTAATAGAGTTAGGCAGGGATAC  
TCACCATTATCGTTTCAGACCCACCTCCCAACCCCGAAGGGA  
CCCGACAGGCCAGAAGGAATCGAAGAAGAAGGTGGAGAGAGA  
GACAGAGGCAGCTCCACTCGATTAGTGCACGGATTCTTAGCA  
CTTTTCTGGGACGACCTGAGGAGTCTGTGCCTCTTCAGCTAC  
CACCACCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTG  
GAACTTCTGGGACGCAGGGGATGGGAAGCCCTCAAATACTGG  
TGGAATCTCCTGCAGTATTGGAGGCAGGAATAACAGAAGAGT  
GCTGTTAGCTTGTTCAATGGCACGGCCATAGCAGTAGCTGAG  
GGGACAGATAGAGTTATAGAAGCTTTACGAAGGGCTTATAGA  
GCTATTCTCCACATACTAGAGAATAAGACAGGGGCTTAGAA  
AGGGCTTTGCTATAAAATGGGTGGCAAGTGGTCAGAAAGTAG  
TGTGGTTAGAAGGCATGTACCTTTAAGACAAGGCAGCTATAG  
ATCTTAGCCGCTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGC  
TAATTCACCTCACGGAAAAGACCAGTTGAACCAGAAGAAGATA  
GAAGAGGCCATGAAGAAGAAAACAACAGATTGTTCTGCTTGC  
TCAGCTGGGGACTTTCCAGAAGGCGCGGCCTGAGTGACTAAG  
CCCCGTTGGGGACTTTCCGAAGAGGCATGAAGGGACTTTCCA  
AGGCAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCT  
CAGATGCTGCATATAAGCAGCTGCTTTCTGCCTGTACTGGGT  
CTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCT  
AGCTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTT  
GAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTG  
GTATCTAGAGATCCCTCAGACCATTTTAGTCCGTGTGGAAA  
TCTCTAGCA

38/47





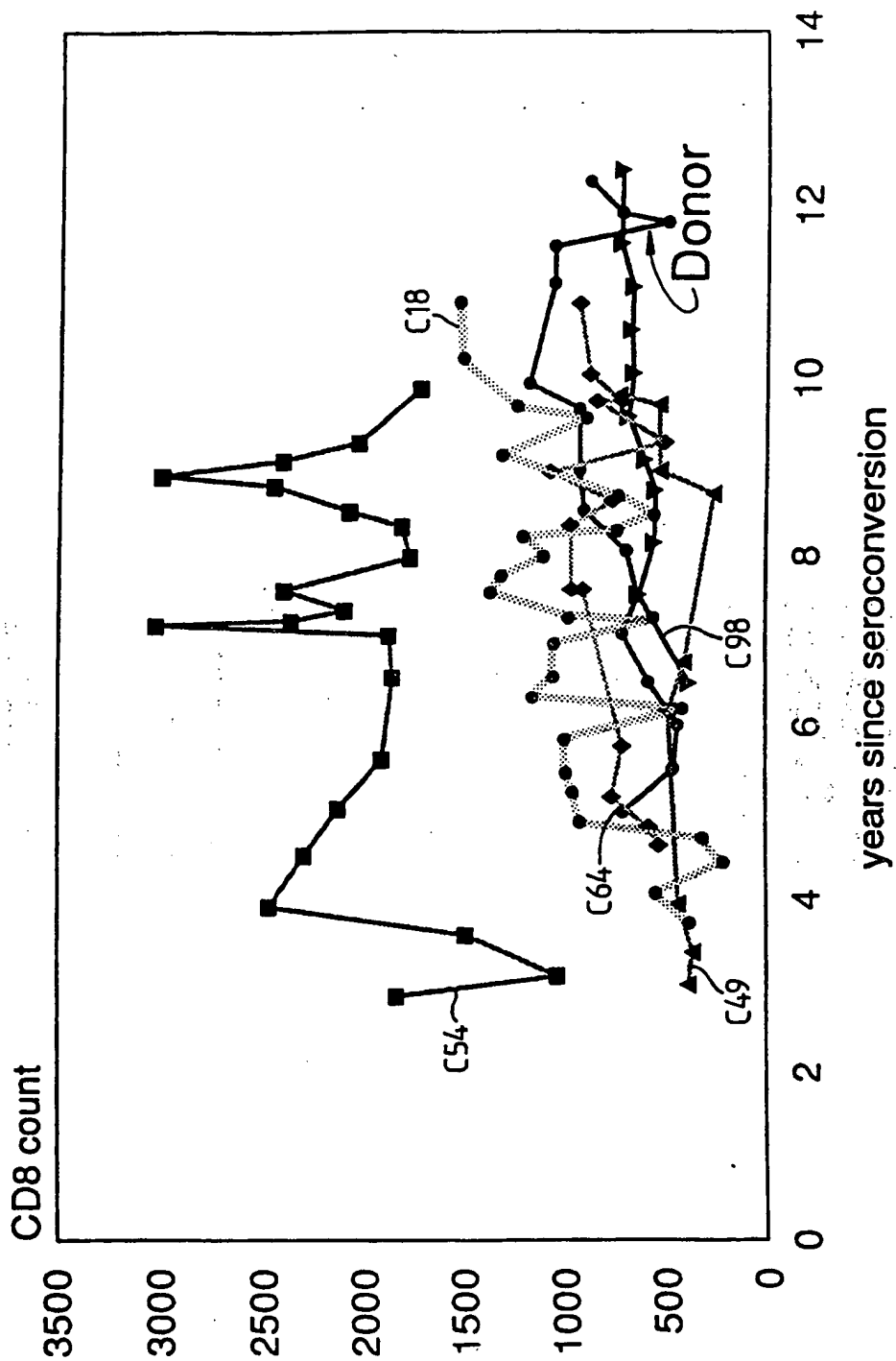
40/47



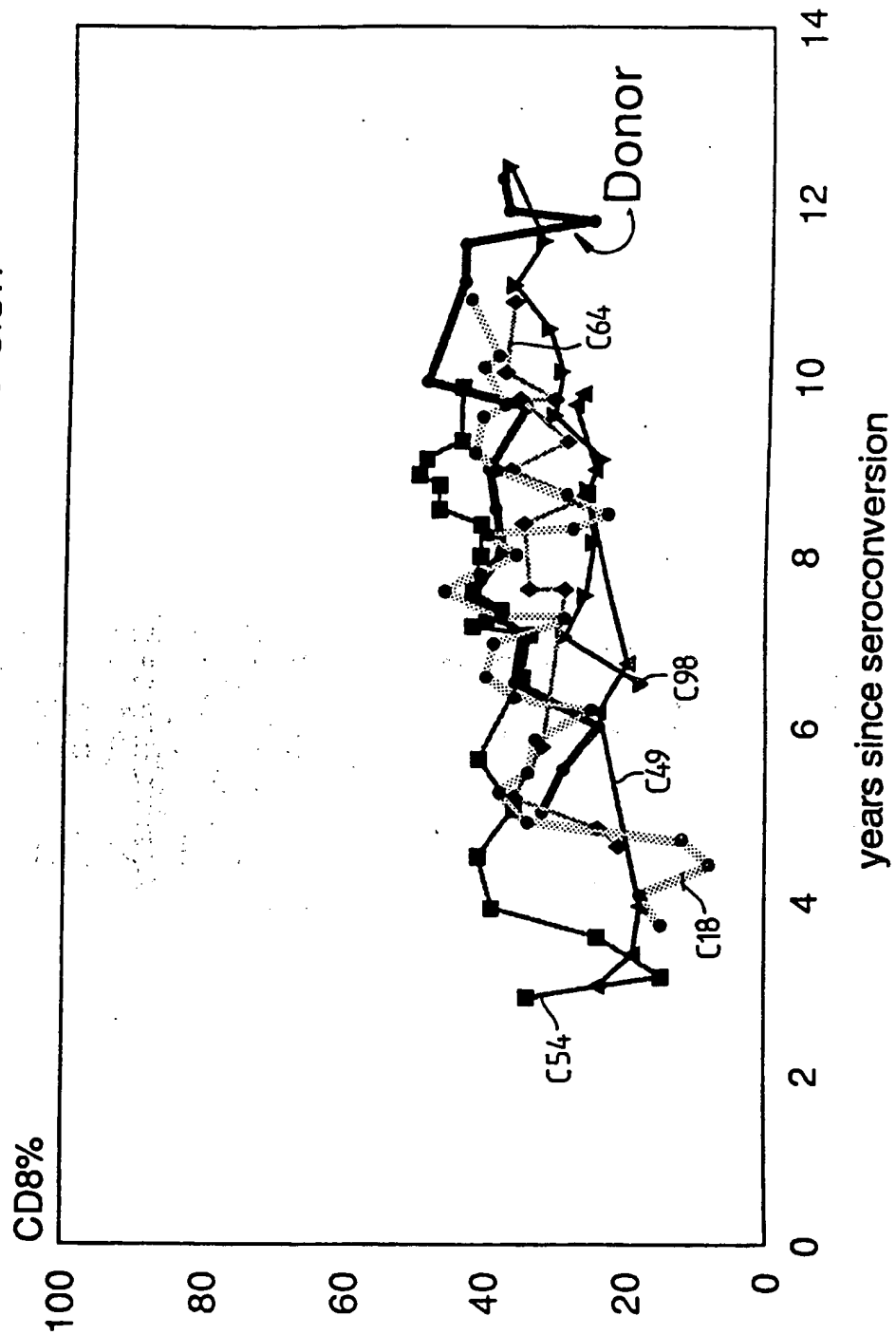


41/47

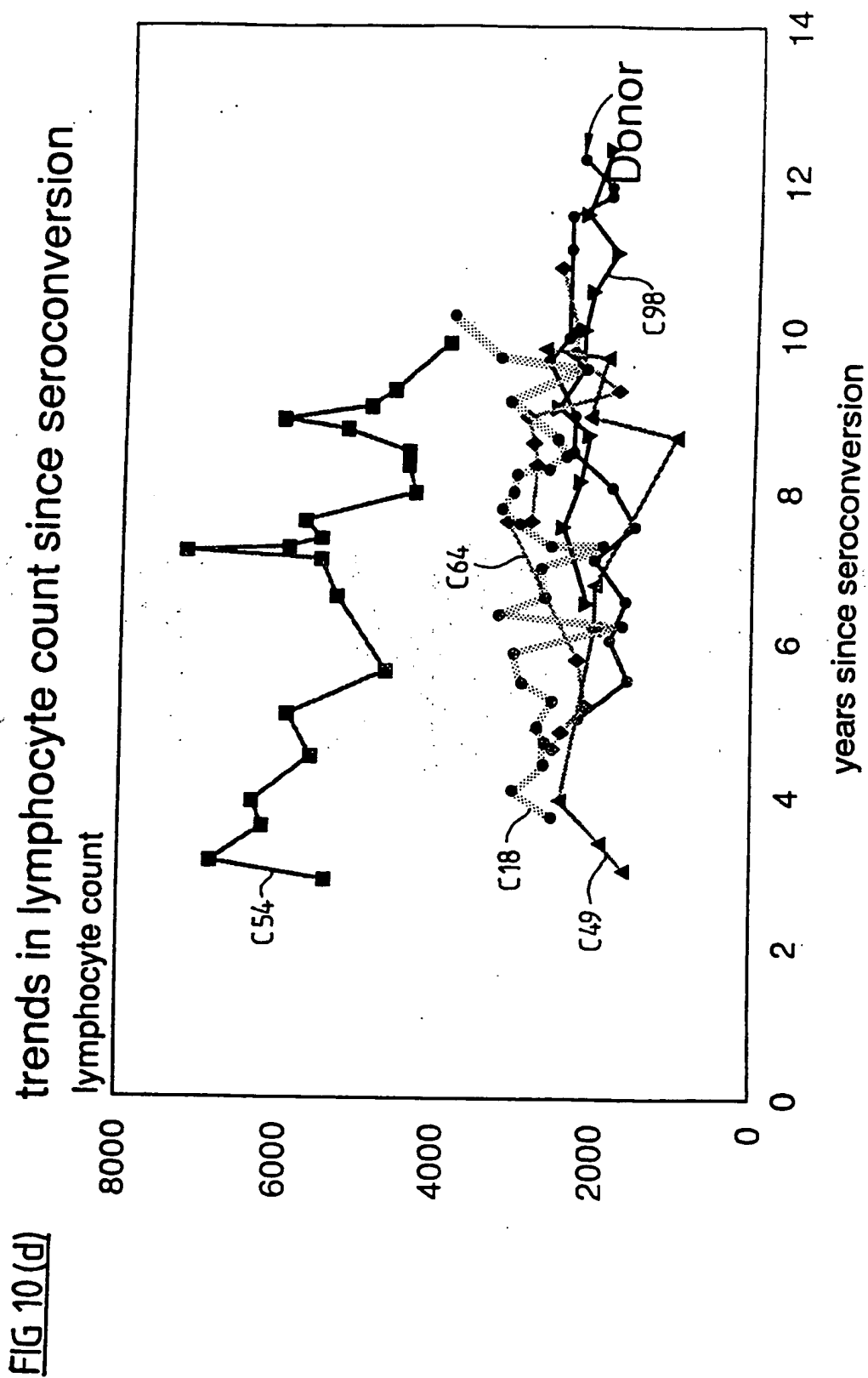
FIG 10(c)(i)  
trends in CD8 since seroconversion



42/47

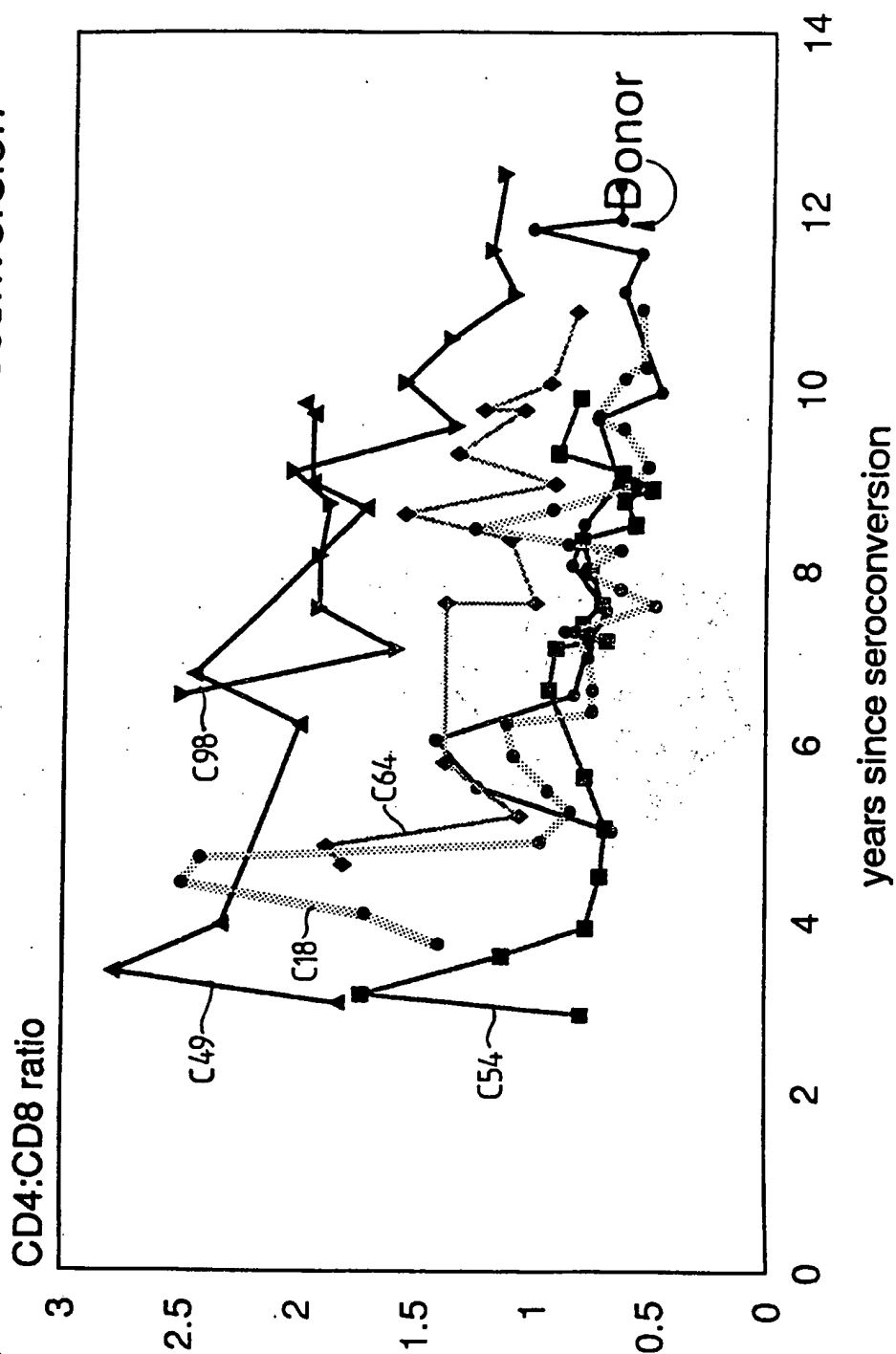
FIG 10 (c)(ii)  
trends in CD8% since seroconversion

43/47



44/47

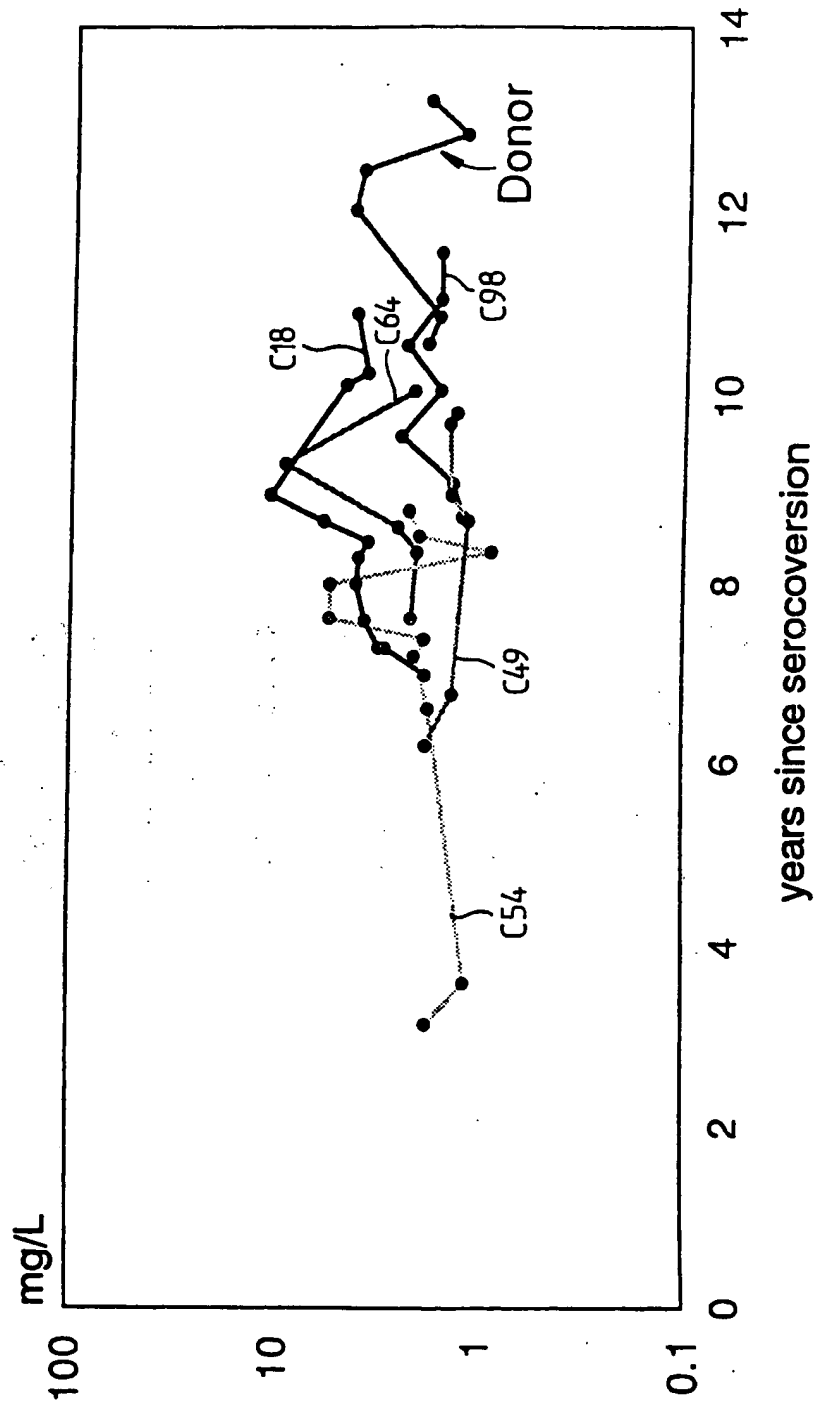
**FIG 10(e)** trends in CD4:CD8 ratio count since seroconversion



45/47

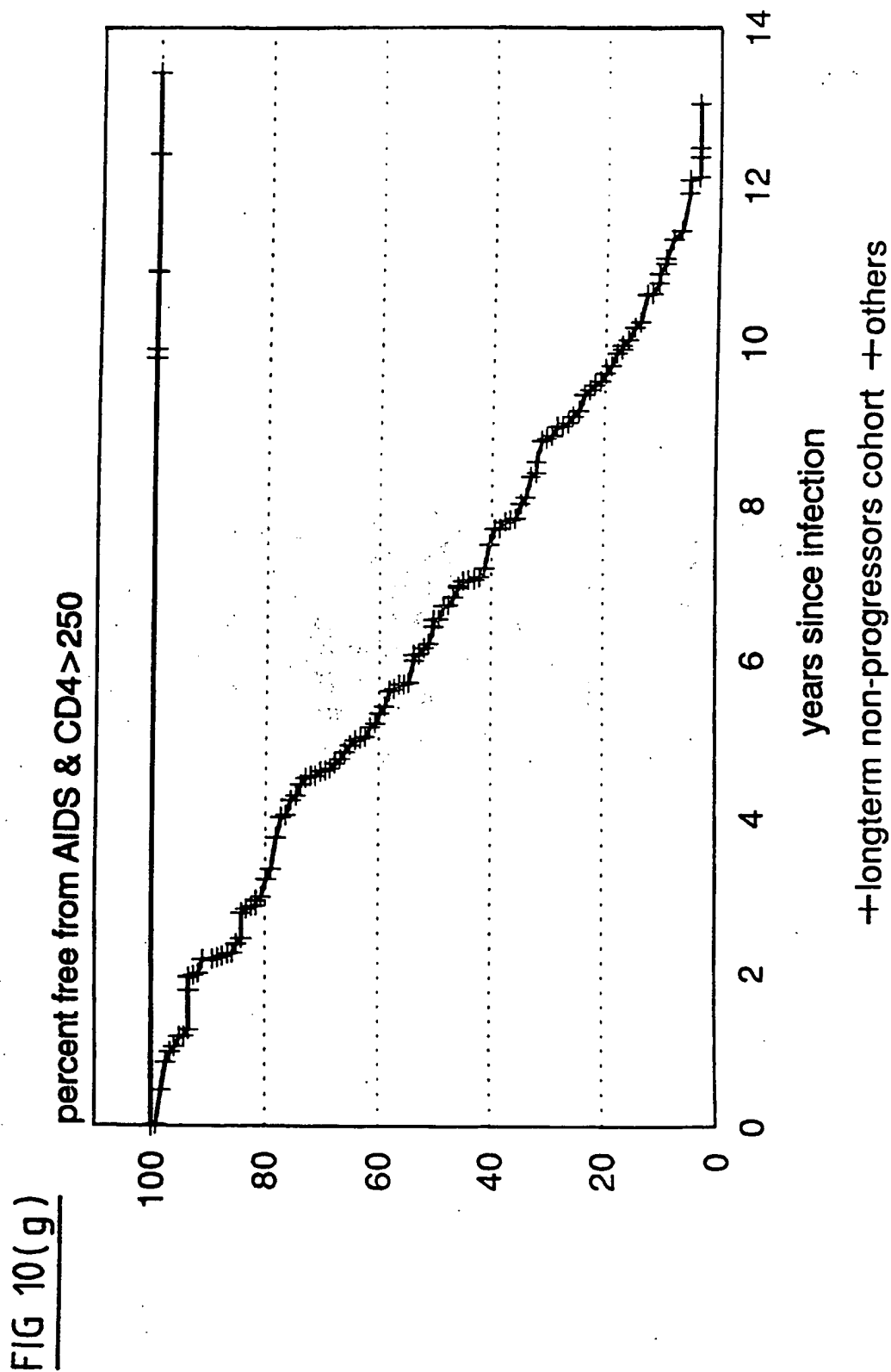
trends in beta-2 microglobulin since seroconversion

FIG 10 (f)



beta-2 microglobulin plotted on a log scale

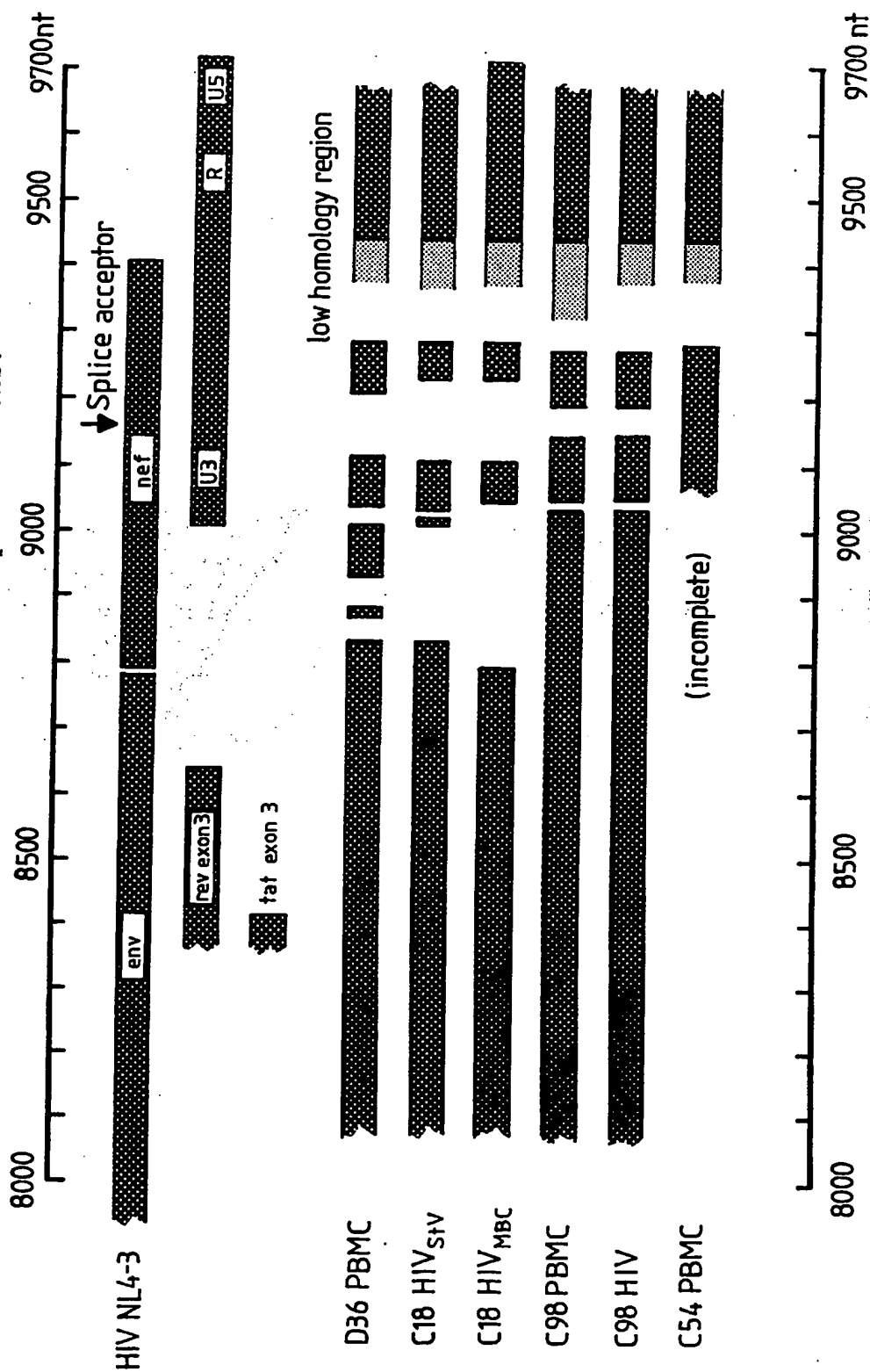
46/47



47/47

**FIG 11**


Diagrammatic sequence alignment against HIV NL4-3 showing deletions in HIV-1 asymptomatic cohort. Deletions shown as gaps in sequence blocks.



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU 95/00063

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> Int. Cl. <sup>6</sup> C12N 7/02, 7/04, 15/48; C12Q 1/68, 1/70; A61K 39/21  According to International Patent Classification (IPC) or to both national classification and IPC				
<b>B. FIELDS SEARCHED</b>  Minimum documentation searched (classification system followed by classification symbols) ELECTRONIC DATABASES AS BELOW  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched AU IPC C12N 7/02, 7/04  Electronic data base consulted during the international search (name of data base, and where practicable, search terms used) DERWENT - WPAT, BIOT: CHEMICAL ABSTRACTS - CASM, KEYWORDS: HIV, LAV, ARV, HTLV, AVIRULENT, ATTENUATED, NON PATHOGENIC, DELETION, NEF, LTR				
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>				
<b>Category *</b>	<b>Citation of document, with indication, where appropriate, of the relevant passages</b>	<b>Relevant to Claim No.</b>		
P, X	WO,A, 94/17825 (UNIVERSITY OF CALIFORNIA) 18 August 1994, see entire document.	1-92		
X	WO,A, 91/19795 (IMMUVAX) 26 December 1991, see entire document.	1-92		
X	WO,A, 92/00987 (HARVARD COLLEGE) 23 January 1992, see entire document.	1-92		
X	WO,A, 92/05864 (CONNAUGHT LABORATORIES LTD) 2 May 1991 see pages 8-9, figures 1-2, claims.	1-92		
<div style="display: flex; justify-content: space-between;"> <div> <input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.         </div> <div> <input checked="" type="checkbox"/> See patent family annex.         </div> </div>				
<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top;"> <b>* Special categories of cited documents :</b>  <b>"A"</b> document defining the general state of the art which is not considered to be of particular relevance  <b>"E"</b> earlier document but published on or after the international filing date  <b>"L"</b> document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  <b>"O"</b> document referring to an oral disclosure, use, exhibition or other means  <b>"P"</b> document published prior to the international filing date but later than the priority date claimed             </td> <td style="width: 50%; vertical-align: top;"> <b>"T"</b> later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  <b>"X"</b> document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  <b>"Y"</b> document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art  <b>"&amp;"</b> document member of the same patent family             </td> </tr> </table>			<b>* Special categories of cited documents :</b> <b>"A"</b> document defining the general state of the art which is not considered to be of particular relevance <b>"E"</b> earlier document but published on or after the international filing date <b>"L"</b> document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) <b>"O"</b> document referring to an oral disclosure, use, exhibition or other means <b>"P"</b> document published prior to the international filing date but later than the priority date claimed	<b>"T"</b> later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention <b>"X"</b> document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone <b>"Y"</b> document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art <b>"&amp;"</b> document member of the same patent family
<b>* Special categories of cited documents :</b> <b>"A"</b> document defining the general state of the art which is not considered to be of particular relevance <b>"E"</b> earlier document but published on or after the international filing date <b>"L"</b> document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) <b>"O"</b> document referring to an oral disclosure, use, exhibition or other means <b>"P"</b> document published prior to the international filing date but later than the priority date claimed	<b>"T"</b> later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention <b>"X"</b> document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone <b>"Y"</b> document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art <b>"&amp;"</b> document member of the same patent family			
Date of the actual completion of the international search 1 May 1995	Date of mailing of the international search report 3 May 1995 (03.05.95)			
Name and mailing address of the ISA/AU  AUSTRALIAN INDUSTRIAL PROPERTY ORGANISATION PO BOX 200 WODEN ACT 2606 AUSTRALIA  Facsimile No. 06 2853929	Authorized officer   ROSS OSBORNE  Telephone No. (06) 2832404			



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU 95/00063

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate of the relevant passages	Relevant to Claim No.
X	WO,A, 90/13641 (SLOAN KETTERING INSTITUTE FOR CANCER RESEARCH) 15 November 1990.	68
Y		70-71
X	The Lancet, Volume 340, issued October 10, 1992, J. Learmont et al, "Long term symptom less HIV-1 infection in recipients of blood products from a single donor". pages 863-867 see entire article.	1-92
Y	Cell, Volume 65, issued May 17, 1991, H.W. Kestler et al "Importance of the nef gene for maintenance of high virus loads and for development of AIDS". pages 651-662. See page 59 column 2 line 54 to page 660 column 1 line 3.	5-19,24-36,40-69,71-92
Y	Science, Volume 258, issued 18 December 1992, M.D. Daniel et al "Protective effects of a live attenuated SIV vaccine with a deletion in the nef gene". pages 1938-1941.	5-19,24-36,40-69,71-92
X	Derwent WPAT Online Abstract Accession Number 93-146253 JP,A, 5078386 (SANYO KOKUSAKU PULP CO) 30 March 1993.	68
X	US 5221610 (INSTITUT PASTEUR) 22 June 1993, see column 5 lines 26-43 and column 16 line 39 to column 20 claim 5.	65-69
X	AU-B-73582/87 (588462) (U.S. DEPARTMENT OF COMMERCE) 9 November 1987	1-4,20-23,37-39,70
Y	whole document.	5-19,24-36,40-69,71-92
Y	Proc. Natl. Acad. Sci. USA. Volume 89 issued November 1992, J.O. Ojwang et al "Inhibition of human immunodeficiency virus type 1 expression by a hairpin ribozyme". pages 10802-10806, see entire article.	68,70-71
P,X	WO,A, 94/29437 (UNIVERSITY OF MEDICINE AND DENTISTRY OF NEW JERSEY) 22 December 1994.	1-7,20-26,37-39,70

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU 95/00063

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international search report has not established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claim Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH

Information on patent family members

**PCT/AU 95/00063**

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report				Patent Family Member			
WO	9119795	CA	2085897	EP	537247		
WO	9417825	AU	58487/94	WO	9417825		
WO	9200987	EP	491930	JP	5501654		
WO	9205864	AU	86591/91	CA	2092553	EP	478842
		EP	479187	FI	931506	JP	6504941
		NO	930993	PT	99148		
WO	9013641	EP	471796	JP	4505261		
AU	87/73587	CN	87103817	DK	288/88	EP	247002
		FI	880275	FR	2599208	NO	880281
		PT	84927	US	4945289	WO	8707469
		YU	923/87				
END OF ANNEX							